FINAL REPORT

Cost-Effective and High-Resolution Subsurface Characterization Using Hydraulic Tomography

ESTCP Project ER-201212

AUGUST 2017

Dr. Chin Man Mok
Walter A. Illman
Tian-Chyi J. Yeh
GSI Environmental Inc.

ESTCP

Page Intentionally Left Blank

Page Intentionally Left Blank

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
15-06-2017	Final Report	

4. TITLE AND SUBTITLE

COST-EFFECTIVE AND HIGH-RESOLUTION SUBSURFACE CHARACTERIZATION USING HYDRAULIC TOMOGRAPHY

5a. CONTRACT NUMBER
W912HQ-12-C-0024

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mok, Chin Man W.; Yeh, T.C.-Jim; Illman, Walter A.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

GSI Environmental Inc.
155 Grand Avenue, 7th Floor
Oakland, CA 94612

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

SERDP/ESTCP
4800 Mark Center Drive
Suite 17D08
Alexandria, VA 22350-6386

10. SPONSOR/MONITOR'S ACRONYM(S)
ESTCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
ER201212

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Hydraulic Tomography (HT) is a high-resolution technology for delineating the spatial distribution of hydraulic conductivity (K) and specific storage (S_s) parameters, which critically affect the performance of subsurface remedial actions at environmental sites. The good technical performance and cost-effectiveness of HT have been demonstrated in this project at two field sites: (1) a local-scale site at a spatial resolution critical to typical source zone remedial actions and (2) a field-scale site typical of Department of Defense (DoD) environmental sites. This project: 1) demonstrated that HT is superior to conventional methods for estimating the spatial distribution of hydrogeologic properties; 2) illustrated that an HT survey can be readily conducted at DoD sites using existing networks of groundwater extraction/injection and observation wells; and 3) developed guidance for HT field implementation and compare costs associated with HT and conventional methods.

15. SUBJECT TERMS Continuous multichannel tubing, Hydraulic Tomography, Hydraulic conductivity, Pumping well, Steady state Hydraulic Tomography, Transient Hydraulic Tomography, Transmissivity

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Chin Man W. Mok
Final Report				117	**19b. TELEPHONE NUMBER** *(include area code)* 510-316-8445

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Page Intentionally Left Blank

FINAL REPORT
Project: ER-201212

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACRONYMS AND ABBREVIATIONS

1-D	One dimensional
2-D	Two dimensional
3-D	Three dimensional
ADEQ	Arizona Department of Environmental Quality
AFP44	Air Force Plant No. 44
a.k.a.	Also known as
ALRS	Apache Leap Research Site
bgs	Below ground surface
BHRS	Boise Hydrogeophysical Research Site
CMT	Continuous multichannel tubing
DCE	Dichloroethene
d	Day
DoD	Department of Defense
EPA	Environmental Protection Agency
Eq(s).	Equation(s)
ft	feet
FR	Final report
HT	Hydraulic Tomography
hr	Hour
K	Hydraulic conductivity
L/min	Liters per minute
ln	Natural logarithm
LZ	Lower Zone
m	Meters
m/s	Meters per second
m/d	Meters per day
m^3/d	Cubic meters per day
MAP	Maximum a posteriori
mbgs	meters below ground surface
NA	Not available
NCRS	North Campus Research Site
pdf	Probability density function

psig	Pounds per square inch gauge
PW	Pumping well
RF	Random field
RMS	Root-mean-square
RPM	Remedial Project Manager
RV	Random variable
SimSLE	Simultaneous Successive Linear Estimator
SLE	Successive Linear Estimator
SSLE	Sequential Successive Linear Estimator
SSHT	Steady state Hydraulic Tomography
S_s	Specific storage
THT	Transient Hydraulic Tomography
T	Transmissivity
TCE	Trichloroethylene
UW	University of Waterloo
UZ	Upper Zone
UZLU	Upper Zone Lower Unit
UZUU	Upper Zone Upper Unit

ACKNOWLEDGEMENTS

The following individuals and organizations contributed to this demonstration project:

- Chin Man W. Mok (PI), GSI Environmental Inc.
- Tian-Chyi Jim Yeh (co-PI), University of Arizona
- Walter A. Illman (co-PI), University of Waterloo
- Zhanfeng Zhao, former student at University of Waterloo
- Yuanyuan Zha, former researcher at University of Arizona
- Michael C.-H. Tso, former student at University of Arizona
- Barbara A. Rudnick, GSI Environmental Inc.
- Yu-Li Wang, University of Arizona
- Richard Balmes, AECOM/URS
- Yemia Hashimoto, formerly with AMEC
- Stuart Neiman, formerly with AGEISS
- Kelly Reis, formerly with AECOM/URS

Page Intentionally Left Blank

EXECUTIVE SUMMARY

OBJECTIVE OF THE DEMONSTRATION

In many cases, especially at recalcitrant sites with complex hydrogeology, inaccurate or inadequate delineation of groundwater flow fields at appropriate resolution has resulted in poor remediation performance. Hydraulic conductivity (K) and specific storage (S_s) are the major parameters governing the fate and transport of contaminants in the subsurface. High-K zones and fractures are fast-flow conduits where transport of dissolved contaminants potentially poses significant threats to downgradient receptors. Low-K zones are potential repositories of contaminant mass that slowly release contaminants and contribute to long-term risks and liability. The overall objective of this project is to provide the Department of Defense (DoD) and its remediation contractors with the Hydraulic Tomography (HT) technology for delineating the spatial distribution of the K and S_s parameters in high resolution. Specific technical objectives are to: 1) demonstrate that HT is superior to conventional methods for estimating the spatial distribution of hydrogeologic properties; 2) illustrate that an HT survey can be readily conducted at DoD sites using existing networks of groundwater extraction/injection and observation wells; and 3) develop guidance for HT field implementation and compare costs associated with HT and conventional methods.

TECHNOLOGY DESCRIPTION

HT concept is analogous to the Computerized Tomography (CT) scanning technology, which is based on combining a series of X-ray images taken from many different angles to make detailed pictures of the physiological structures inside a human body. HT involves sequentially conducting a series of aquifer hydraulic tests (HT survey). The hydraulic stresses in the subsurface are perturbed differently in each test, and the resulting potentiometric head changes over a well network are monitored. Each test is comparable to taking a snapshot of the aquifer heterogeneity, and the whole HT survey is analogous to hydraulically scanning the subsurface. The complete data set of observed potentiometric head responses at multiple locations are jointly analyzed through a consistent mathematical model, which provides detailed spatial distribution of hydraulic properties of the aquifer, patterns of connectivity of highly conductive zones, locations of low conductive zones, and the uncertainties associated with the spatial distribution (HT analysis).

DEMONSTRATION RESULTS

The technical performance and cost-effectiveness of HT have been demonstrated at two field sites: (1) the University of Waterloo (UW) North Campus Research Site (NCRS), which is a local-scale site extensively instrumented at a spatial resolution critical to typical source zone remedial actions, and (2) the Air Force Plant No. 44 (AFP44) site, which is at a field-scale typical of Department of Defense (DoD) environmental sites with an existing pump-and-treat system and monitoring well network.

The results from the demonstrations at both sites confirmed that the HT is more accurate than conventional site characterization techniques. The results clearly indicate that the HT predictions of hydraulic responses during pumping tests outperform those of conventional models.

The HT predictions are unbiased and have smaller root-mean-square of prediction errors. The results also confirmed that the HT results are less uncertain than the results from conventional site characterization techniques. The HT results are consistent with the current knowledge of the spatial distribution of the more permeable and less permeable regions. The demonstration at the AFP44 site illustrated that HT is particularly cost-effective and can readily be applied at sites with existing well networks and pump-and-treat systems. The only costs for conducting the HT site characterization were the labor costs for conducting the pumping tests and performing the HT model inversion. HT is a "user-friendly" site characterization technology. The skills and equipment needed for conducting HT survey are the same as those commonly used in conventional site characterization. The input data required for model inversion by HT are the same as the data used in groundwater model development and calibration, such as the input data for parameter estimation using the commonly used software PEST and MODFLOW. The demonstration results illustrated that HT is able to delineate low-K zones consistent with the available lithologic data locally. In addition, it can infer the hydraulic continuity of the low-K regimes in between available lithologic information. It provides information as to whether these regimes are hydraulically functioning as competent barriers. In conjunction with available chemical concentration data, the information is useful for evaluating potential residual sources.

IMPLEMENTATION ISSUES

If the on-site water treatment system is not available or not suitable for the extracted water from an HT survey, temporary storage and transportation options should be discussed, with consideration of the pumping rates and durations required for showing sufficient drawdown responses. If injection tests are required for the HT survey, a suitable source of injection water, such as clean or treated water, needs to be found and its transportation planned accordingly.

If additional wells are needed, and especially if they need to be installed in areas with high chemical concentration, pertinent regulatory approval and permits might be required. This is a similar issue with conventional well installation. If the HT pumping tests involve groundwater extraction, pumping permits might be required. In addition, permits for the discharge to the on-site or off-site treatment systems need to be acquired. Depending on the application process, extraction water sampling might be necessary. Similarly, permits might have to be obtained for the water injection, with a potential sampling of the injection water.

The key factors to be considered in making a decision as to whether HT is appropriate for a site include cost-effectiveness, timing and duration, knowledge of background hydraulic stresses, and chemical mobilization. The cost-effectiveness depends on the appropriate number of wells, which is dictated by the spatial resolution needed to meet the objectives and whether existing wells and treatment system are adequate. If an existing well and treatment system can be utilized, the costs associated with HT is minimal.

In addition, water level changes due to HT pumping tests might cause chemicals to move during the tests. The duration of the pumping tests is usually short, and the amount of the associated chemical movement is typically small. However, if the aquifer is very permeable, a large pumping rate might be required to generate a measurable hydraulic response signal. On the other hand, if the aquifer is relatively impermeable, the well yield might be small, and a longer HT pumping test duration might be needed.

1.0 INTRODUCTION

Hydraulic Tomography (HT) is a high-resolution subsurface characterization technology for delineating the spatial distribution of hydraulic conductivity (K) and specific storage (S_s) parameters. These parameters are the major factors governing the fate and transport of contaminants in the subsurface, thus critically affecting the performance of remedial actions at environmental sites. The technical performance and cost-effectiveness of HT have been demonstrated in this project at two field sites. The first demonstration was performed at the North Campus Research Site (NCRS) at the University of Waterloo (UW) in Canada. It is a local-scale site extensively instrumented at a spatial resolution critical to typical source zone remedial actions. The second demonstration was conducted at the Air Force Plant No. 44 (AFP44) site in Tucson, Arizona. It is at a field-scale typical of Department of Defense (DoD) environmental sites. This site has an existing pump-and-treat system and monitoring well network. This Final Report (FR) is a comprehensive technical document summarizing the project's activities, results, and conclusions.

1.1 BACKGROUND

The groundwater flow field is a critical factor dictating the fate and transport of contaminants in the subsurface. It is highly dynamic and heterogeneous. Contaminant dispersion is a result of spatial groundwater flow variation. More importantly, high-K zones and fractures, such as buried gravelly stream channels, are fast-flow conduits where transport of dissolved contaminants occurs. These preferential pathways potentially pose significant threats to downgradient receptors. Low-K zones, such as clayey lenses, are potential repositories of contaminant mass that slowly releases contaminants due to back-diffusion. These residual sources contribute to long-term environmental risks and liability.

In many cases, especially at recalcitrant sites with complex hydrogeology, inaccurate or inadequate delineation of groundwater flow fields at appropriate resolution has resulted in poor remediation performance. Examples of this, just to name a few, are a pump-and-treat system fails to cost-effectively or efficiently contain contaminated groundwater; a chemical of concern migrates downgradient along unidentified pathways; an injected substrate does not reach targeted treatment zones or has insufficient residence time to enhance bioremediation; an impermeable barrier does not fully intercept contaminant migration pathways; or, a monitoring well network is not installed at appropriate locations to collect useful information. Many pump-and-treat sites in the United States have been operated for more than fifteen years without achieving remediation goals. Operating and maintaining these systems is often costly. Many of them are now undergoing optimization and re-evaluation. Despite the best remediation efforts, sites with complex heterogeneous hydrogeology continue to act as long-term environmental liabilities. Some of these sites are even in the process of considering a technical impracticability waiver application. Accurately depicting the subsurface hydrogeology in both contaminant source zones and dissolved plume areas is crucial for reliable assessments of potential risks to nearby receptors and design of effective remediation systems. Therefore, subsurface characterization techniques that provide high-resolution information are critical for improving performance of existing systems and/or for developing alternative remedial action to achieve groundwater cleanup goals in a reasonable timeframe.

Conventional hydrogeological characterization techniques, such as borehole core or cuttings samples, generally provide local-scale geologic, lithologic, and/or hydrostratigraphic data at a few locations. Spatially interpolating or extrapolating this punctual information across the area of concern is subjective. In addition, this information does not directly provide hydraulic parameter values. Estimating the spatial distribution of K and S_s parameters based on this information is inherently uncertain. Although high-resolution information may be obtained using borehole sampling, it is invasive and cost-intensive, especially in deep formations.

Aquifer tests may be performed at a site. The results are commonly analyzed to estimate K and S_s-values using analytical solutions based on the simplified assumption that the aquifer is homogeneous and uniform (e.g., Theis' or Cooper-Jacob method). Such analyses yield equivalent properties that somewhat represent the typical properties between the pumping well and monitoring well within the cone of depression. Geophysical methods have increasingly been used to supplement conventional characterization by producing a high-resolution image of the subsurface. Although these methods can be relatively quick and inexpensive to perform, they only provide a high-resolution image of geophysical properties instead of hydrogeologic properties. Site-specific petrophysical relationships may have to be developed to translate the geophysical properties to corresponding hydrogeologic properties, leading to considerable uncertainty.

1.2 OBJECTIVE OF THE DEMONSTRATION

The overall objective of this project is to provide the DoD and its remediation contractors with the HT technology for delineating the spatial distribution of the K and S_s parameters in high resolution. Specific technical objectives are to: 1) demonstrate that HT is superior to conventional methods for estimating the spatial distribution of hydrogeologic properties; 2) illustrate that an HT survey can be readily conducted at DoD sites using existing networks of groundwater extraction/injection and observation wells; and 3) develop guidance for HT field implementation and compare costs associated with HT and conventional methods.

1.3 REGULATORY DRIVERS

Regulations protecting water resources require environmentally impaired aquifers to be remediated to an acceptable condition. Sources and impacted zones might need to be contained to prevent further expansion of the contamination extent. The success of remedial action at a site in achieving clean-up goals, as well as the ability of a containment system to control contaminant migration, hinges upon whether groundwater flow field can be adequately delineated. HT is a technology for depicting the groundwater flow field in high resolution. Incorporating the results from HT in remediation and containment operations would increase the reliability of remedial action and the chance of meeting regulatory requirements.

2.0 TECHNOLOGY

HT is a new generation of hydraulic testing and analysis technology used to image the spatial distribution of the subsurface K and S_s parameters in high-resolution (K and S_s tomograms). The development of HT has been funded by SERDP over the past decade. HT has been validated in numerical experiments, controlled laboratory experiments, and field experiments.

2.1 TECHNOLOGY DESCRIPTION

The HT concept is comparable to a person viewing an object from different angles to gain more details of the geometry of an object. An example of this analogous concept employed in medical sciences is the Computerized Tomography (CT) scanning technology, which is based on combining a series of X-ray images taken from many different angles to make detailed pictures of the physiological structures inside a human body.

HT involves sequentially conducting a series of aquifer hydraulic tests (HT survey). The hydraulic stresses in the subsurface are perturbed differently in each test, and the resulting potentiometric head changes over a well network are monitored. Each test is comparable to taking a snapshot of the aquifer heterogeneity, and the whole HT survey is analogous to hydraulically scanning the subsurface. The complete data set of observed potentiometric head responses at multiple locations are jointly analyzed through a consistent mathematical model, which provides detailed spatial distribution of hydraulic properties of the aquifer, patterns of connectivity of highly conductive zones, locations of low conductive zones, and the uncertainties associated with the spatial distribution (HT analysis). The HT technology is schematically illustrated in Figure 2-1.

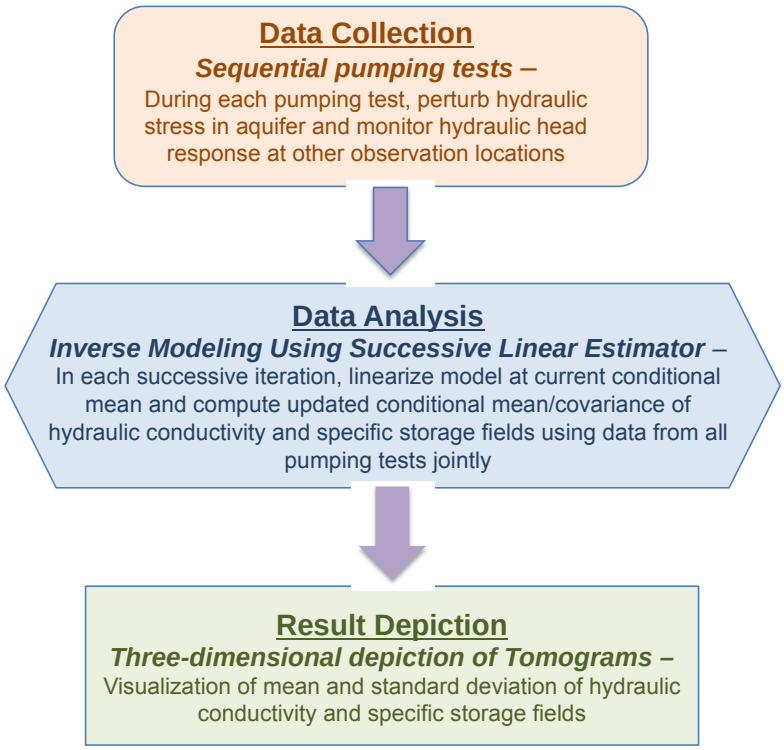


Figure 2-1. Hydraulic Tomography Concept

The novelty of the HT technology demonstrated in this study is the collection of non-redundant hydraulic information from different pumping tests in HT survey and the inclusion of all data in HT analysis without making a presumption of the form of spatial K and S_s distributions. It is different from the zonation and pilot point approaches that are commonly utilized to represent the spatial K and Ss distributions subjectively. Figure **2-2** shows an example of a three-dimensional distribution of a K-field delineated by HT. The resolution of HT results depends upon the spacing of wells.

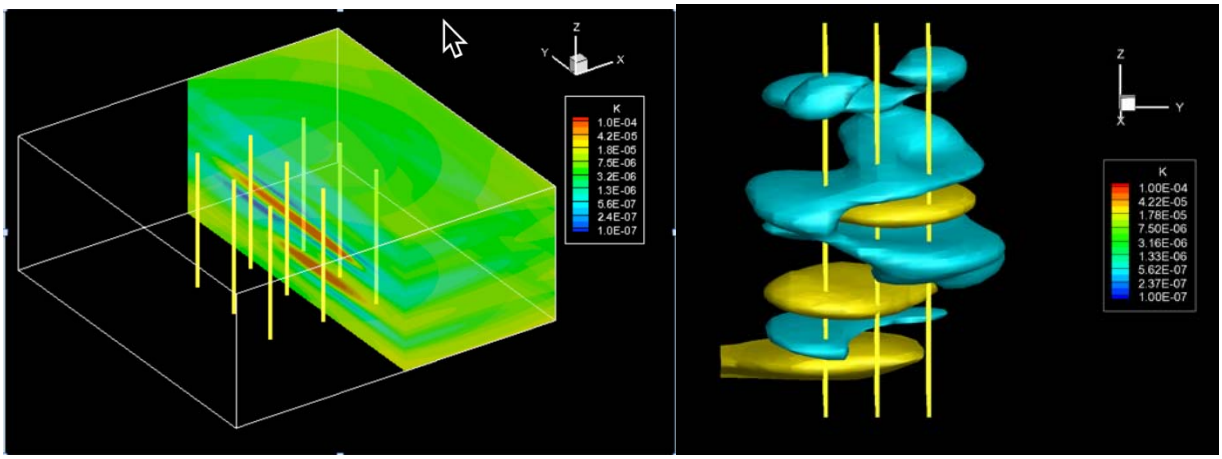


Figure 2-2. Three-Dimensional Distribution of K Parameters

2.1.1 HT Survey

Hydraulic stresses are commonly perturbed in an aquifer test by turning extraction and/or injection well(s) on or off to induce propagation of potentiometric head changes at multiple locations throughout the aquifer. If an aquifer interacts with the surface water regime in the vicinity, such as a river, surface water stage changes during rainfall events naturally generate hydraulic stress perturbations in the aquifer. At sites where ongoing remedial operations include pump-and-treat systems, HT surveys can be conducted by simply modifying the pumping rates or by taking advantage of the pumping shut-off and commencement operations. In operating pump-and-treat sites, shutting down the pump-and-treat system for an extended period of time for characterization may violate the site's record of decision. On the other hand, strong hydraulic responses are generated by the pumping wells. Regrettably, these signals are rarely exploited to improve site characterization. An approach that utilizes ongoing pump-and-treat signals to improve site characterization would be attractive to optimizing remediation strategies throughout its course.

Figure 2-3 shows an example of different potentiometric responses at various monitoring locations in response to HT aquifer tests at three different locations. The greyish bars represent the pumping intervals and the location of the pumping well. The sizes of the purple bubbles are proportional to the magnitudes of the normalized potentiometric head responses. A large bubble is an indication of a stronger hydraulic connection between the pumping and monitoring locations.

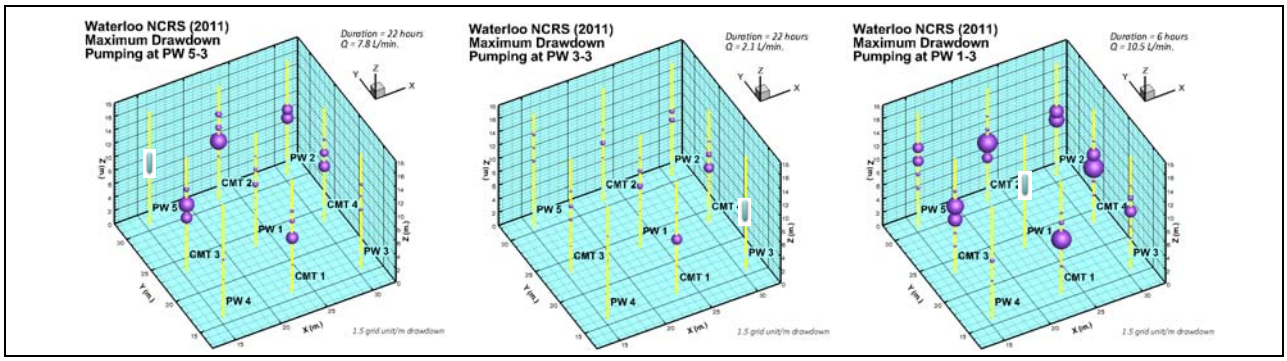


Figure 2-3. Hydraulic Responses to Pumping Tests

2.1.2 HT Analysis

Groundwater flow hydraulic response, as represented by the following governing equation, is dictated by the spatial distribution of the K and S_s parameters.

$$\nabla \cdot \left(K \nabla h \right) + q = S_s \frac{\partial h}{\partial t} \tag{2-1}$$

At any location in the aquifer, the K parameter (similarly, the S_s parameter) is uncertain and has an infinite number of possible values. In HT, the K parameter at a point is treated as a **Random Variable (RV**). We conceptualize the spatially distributed K as a collection of an infinite number of RVs in space, which is referred to as a **Random Field (RF)**. This random field thus has an infinite number of possible spatial distribution patterns. If we also have some samples of K-values at the site, we can further tailor the possible K-fields to the site-specific ones. The RVs at two locations in an aquifer might be correlated. The correlation usually becomes smaller as the separation distance between the two locations increases. A RF model is typically represented in a geostatistical context by (1) probability distribution to characterize the uncertainty of the RV at a point and (2) correlation (or variogram) function to characterize the relationship between correlation and separation distance. HT analysis typically starts with an initial geostatistical model developed using available geologic information. This geostatistical model is referred to as the prior distribution model in a Bayesian statistical framework. HT analysis involves updating the statistical model using the data from HT survey in the Bayesian framework. The resulting K-field is called the conditional effective K-field. In addition, HT analysis estimates the uncertainty associated with the estimated K-field. This variance informs us the likelihood that the estimate K-field can deviate from the true K-field.

The **Successive Linear Estimator (SLE)** used in our HT analysis adopts a highly parameterized heterogeneous conceptual model, which discretizes the 3-D domain of the RF site into N elements. The hydraulic parameter of the N elements (e.g., the natural logarithm of hydraulic conductivity K, ln$K$) composes a **(**N×1**)** vector. The model then considers these hydraulic parameters with prior (unconditional) mean **Y (**N×1**)**, and the prior perturbations **y (**N×1**)**, respectively. These perturbations represent the spatial variability of the parameters.

SLE estimates the most likely parameter value (i.e., conditional effective value) for each element, given (conditioned with) the observed drawdown (or head) data from the HT survey. Suppose during an HT test that we have collected *M* observed heads in time and space, denoted by the data vector **d**. The estimates of parameter fields, given the observation, are $\hat{\mathbf{Y}}_c$ (subscript c denotes conditional), and are iteratively determined using the following linear estimator (Yeh et al., 1996):

$$\hat{\mathbf{Y}}_c^{(r+1)} = \hat{\mathbf{Y}}_c^{(r)} + \boldsymbol{\omega}^{\mathrm{T}} \left(\mathbf{d} - \mathbf{G}\left(\hat{\mathbf{Y}}_c^{(r)} \right) \right) \tag{2-2}$$

where *r* is the iteration index; $\mathbf{G}(\cdot)$ indicates the nonlinear relationship between **Y** and **d** (i.e., a forward groundwater flow model), which produces the simulated heads at the observation locations and times using the parameters obtained at iteration *r*. The coefficient matrix, $\boldsymbol{\omega}$ (*M*×*N*), denotes the weights, which assign the contribution of difference between the observed and simulated head at each observation location and time to a previously estimated parameter value at each element. The superscript T denotes the transpose.

The coefficient matrix $\boldsymbol{\omega}$ is determined by solving the following equations (Yeh et al., 1996):

$$\left[\boldsymbol{\varepsilon}_{dd}^{(r)} + \mathbf{R} \right] \boldsymbol{\omega}^{(r)} = \boldsymbol{\varepsilon}_{dy}^{(r)} \tag{2-3}$$

where **R** is the covariance matrix (*M*×*M*) of the measurement error associated with head measurements. The solution of Eq. (2) requires knowledge of covariance $\boldsymbol{\varepsilon}_{\mathrm{dd}}$ and cross-covariance $\boldsymbol{\varepsilon}_{\mathrm{dy}}$, which can be derived from the first-order numerical approximation (Yeh et al., 1996):

$$\begin{aligned} \boldsymbol{\varepsilon}_{dd}^{(r)} &= \mathbf{J}_d^{(r)} \boldsymbol{\varepsilon}_{yy}^{(r)} \mathbf{J}_d^{(r)\mathrm{T}}, \\ \boldsymbol{\varepsilon}_{dy}^{(r)} &= \mathbf{J}_d^{(r)} \boldsymbol{\varepsilon}_{yy}^{(r)} \end{aligned} \tag{2-4}$$

where $\mathbf{J}_{\mathrm{d}}$ (*M*×*N*) is the sensitivity (Jacobian) matrix of head data with respect to the element-wise parameter using the parameters estimated at current iteration. At the beginning of the iteration (when *r*=0), $\boldsymbol{\varepsilon}_{yy}$ is the unconditional covariance matrix of parameters **Y**, which is traditionally constructed by a prior variance, correlation lengths, and a covariance model (see Section 2.1.3). After that ($r \geq 1$), the residual or conditional covariance function of parameters are updated as (Yeh and Liu, 2000):

$$\boldsymbol{\varepsilon}_{yy}^{(r+1)} = \boldsymbol{\varepsilon}_{yy}^{(r)} - \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{\varepsilon}_{dy} \tag{2-5}$$

The SLE bears the concept of cokriging or stochastic linear estimator equation (e.g., unbiased estimates with minimum variance). The nonlinearity between parameters and heads is dealt with each successive iteration. At iteration *r*=0, SLE requires guessed values for the mean ($\hat{\mathbf{Y}}_c^{(0)}$ used in Eq. (1)) and covariance function ($\boldsymbol{\varepsilon}_{yy}^{(0)}$ used in Eqs. (3-4)). In the view of the Bayesian statistics, $\hat{\mathbf{Y}}_c^{(0)} = \mathbf{Y}_{prior}$ and $\boldsymbol{\varepsilon}_{yy}^{(0)} = \mathbf{C}_{prior}$ are the prior information of the unknown parameter field.

Afterward, SLE updates the mean and the covariance at each iteration due to gradual assimilation of the observation information and reduces the uncertainty of the estimate.

SLE is similar to the maximum a posteriori (MAP) inverse approach, but it is different. As pointed out by Carrera and Glorioso (1991), the cokriging equation produces the same estimate of the first iteration of maximum posterior approaches if the initial guess mean is taken as prior. Let $p(\mathbf{Y}|\mathbf{d})$ be the probability density function (pdf) of model parameter $\mathbf{Y}$ conditioned on the data set $\mathbf{d}$, and $p(\mathbf{Y})$ is the prior pdf. The Bayes theorem gives the pdf of model parameter $\mathbf{Y}$ after the assimilation of the data $\mathbf{d}$ (Fienen et al., 2009):

$$p(\mathbf{Y}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{Y})\,p(\mathbf{Y}) \tag{2-6}$$

If the prior pdf $p(\mathbf{Y})$ can be approximated as Gaussian, with mean $\mathbf{Y}_{\text{prior}}$ ($N\times1$) and covariance $\mathbf{C}_{\text{prior}}$ ($N\times N$), and the error in observation $\mathbf{d}$ ($M\times1$) are normally distributed with zero mean and covariance $\mathbf{R}$ ($M\times M$), Eq. (5) becomes

$$\ln p(\mathbf{Y}|\mathbf{d}) \propto -\frac{1}{2}\left[\left(\mathbf{G}(\mathbf{Y})-\mathbf{d}\right)^{T}\mathbf{R}^{-1}\left(\mathbf{G}(\mathbf{Y})-\mathbf{d}\right)+\left(\mathbf{Y}-\mathbf{Y}_{prior}\right)^{T}\mathbf{C}_{prior}^{-1}\left(\mathbf{Y}-\mathbf{Y}_{prior}\right)\right] \tag{2-7}$$

Using the Gauss-Newton method to minimize the objective function $-\ln p(\mathbf{Y}|\mathbf{d})$, the $(r+1)^{\text{th}}$ iterative estimate of parameter $\mathbf{Y}$ is (Chen and Oliver, 2013):

$$\hat{\mathbf{Y}}_c^{(r+1)} = \hat{\mathbf{Y}}_c^{(r)} + \left(\mathbf{J}_d^{(r)\mathrm{T}}\mathbf{R}^{-1}\mathbf{J}_d^{(r)} + \mathbf{C}_{prior}^{-1}\right)^{-1}\left[\mathbf{J}_d^{(r)\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{d}-\mathbf{G}\left(\hat{\mathbf{Y}}_c^{(r)}\right)\right)+\mathbf{C}_{prior}^{-1}\left(\mathbf{Y}_{prior}-\hat{\mathbf{Y}}_c^{(r)}\right)\right] \tag{2-8}$$

with estimation covariance of

$$\boldsymbol{\varepsilon}_{yy}^{(r+1)} = \left(\mathbf{J}_d^{(r)\mathrm{T}}\mathbf{R}^{-1}\mathbf{J}_d^{(r)} + \mathbf{C}_{prior}^{-1}\right)^{-1} \tag{2-9}$$

Through several linear algebraic manipulations (see Eqs. (1.106-1.107) in (Tarantola, 2005), we have

$$\hat{\mathbf{Y}}_c^{(r+1)} = \hat{\mathbf{Y}}_c^{(r)} + \left(\mathbf{Y}_{prior}-\hat{\mathbf{Y}}_c^{(r)}\right)+\mathbf{C}_{prior}\mathbf{J}_d^{(r)\mathrm{T}}\left(\mathbf{J}_d^{(r)}\mathbf{C}_{prior}\mathbf{J}_d^{(r)\mathrm{T}}+\mathbf{R}\right)^{-1}\left[\left(\mathbf{d}-\mathbf{G}(\mathbf{Y})\right)+\mathbf{J}_d^{(r)}\left(\mathbf{Y}_{prior}-\hat{\mathbf{Y}}_c^{(r)}\right)\right] \tag{2-10}$$

$$\boldsymbol{\varepsilon}_{yy}^{(r+1)} = \mathbf{C}_{prior}-\mathbf{C}_{prior}\mathbf{J}_d^{(r)\mathrm{T}}\left(\mathbf{J}_d^{(r)}\mathbf{C}_{prior}\mathbf{J}_d^{(r)\mathrm{T}}+\mathbf{R}\right)^{-1}\mathbf{J}_d^{(r)}\mathbf{C}_{prior} \tag{2-11}$$

By comparing Eqs. (1-4) and Eqs. (9-10), it is found that the calculated $\hat{\mathbf{Y}}_c^{(r+1)}$ and $\boldsymbol{\varepsilon}_{yy}^{(r+1)}$ have the same forms in SLE and MAP formulations if we set initial guess mean $\hat{\mathbf{Y}}_c^{(0)}$ as the prior $\mathbf{Y}_{\text{prior}}$. In other words, the first iteration of SLE and MAP yield the same estimated mean and covariance. However, after the first iteration, SLE uses the updated mean (e.g., substituting $\mathbf{Y}_{\text{prior}}$ by $\hat{\mathbf{Y}}_c^{(r)}$ in the Eq. (9)) and updated covariance function (e.g., substituting $\mathbf{C}_{\text{prior}}$ by $\boldsymbol{\varepsilon}_{yy}^{(r)}$ in the Eq. (10)) as prior information.

That is, MAP uses static prior information while SLE recursively uses updated mean and covariance from the last iteration as prior (This is similar to Kalman filter in signal analysis, where new observation in time is added). In other words, a posterior mean and covariance at iteration r serve as a prior at iteration r+1. The logic behind this is that the inverse model gradually learns from observation data for every iteration, and updates the probability density function of the uncertainty associated with the estimates at every iteration (Yeh et al., 1996; Yeh and Liu, 2000). Figure 2-4 shows the schematic of the SLE computational steps.

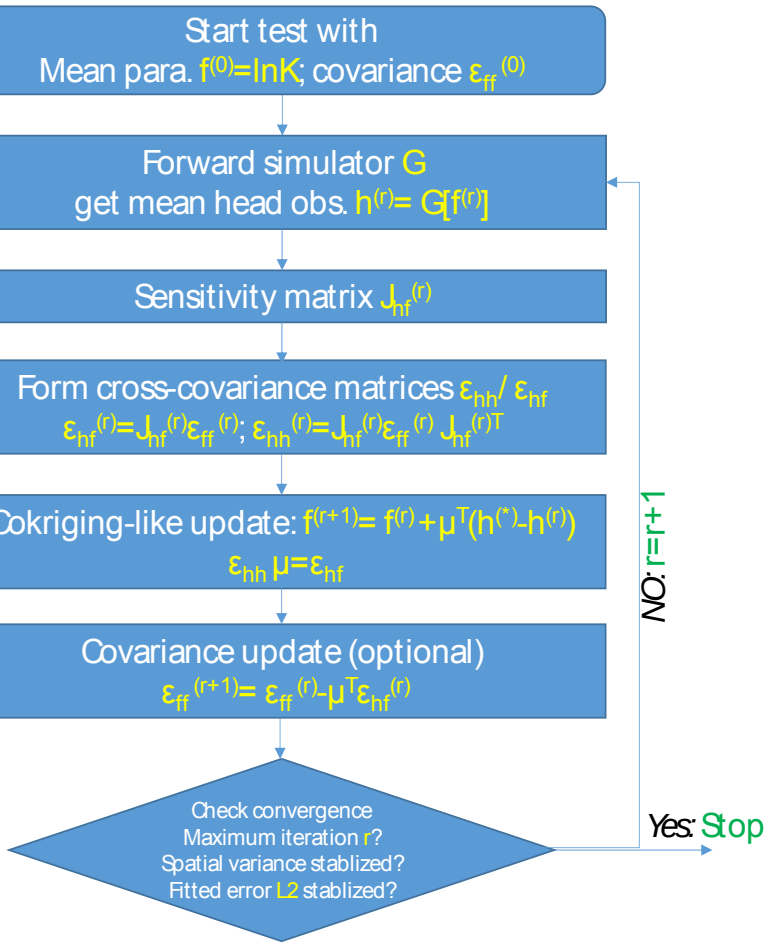


Figure 2-4. Flowchart Summarizing the Successive Linear Estimator (SLE) Methodology

Only parameter K is considered in the example.

In addition, it is noteworthy that the SLE is different from the well-known pilot point method. For reducing computational cost and making the inverse problem well-defined (Yeh et al., 2015), the pilot point method uses only a few selected pilot points, where hydraulic parameter values are estimated by a nonlinear algorithm minimizing the simulated and observed hydraulic head differences. The entire parameter field is obtained afterward by kriging based on the parameter values at the pilot point locations and the unconditional covariance function of the parameter (McLaughlin and Townley, 1996); Soueid Ahmed et al., 2015). That is, the final parameter field and the estimated parameters at pilot points are not linked by the governing flow equation, and the result thus could be suboptimal (Huang et al., 2011).

2.2 TECHNOLOGY DEVELOPMENT

The novelty of the HT technology demonstrated in this study is the strategic collection of non-redundant hydraulic information by multiple pumping tests and the inclusion of all the data in HT analysis without using an under-parameterized model based on the subjective presumption of the spatial form of K and S_s zonation or pilot points. Huang et al. (2011) discussed the limitation of using the pilot point method for tomographic surveys.

Neuman et al. (1984) was the one of the first studies involving the use of data from multiple pumping tests. They considered pumping from three wells alternatively and derived three equations to solve for three unknown anisotropic components of the effective transmissivity of an equivalent homogeneous aquifer. Hsieh et al. (1985) applied the same concept to estimate the anisotropic effective hydraulic conductivity tensor of equivalent homogeneous fractured rocks. However, these studies did not consider aquifer heterogeneity.

In 1995 and 1996, Sandia National Laboratories installed seven boreholes over an area of 50 × 50 m in the Culebra dolomite of the Rustler Formation within southeastern New Mexico's Delaware Basin. The Culebra is known as a laminated to thinly bedded argillaceous dolomite with abundant open and gypsum-filled fractures. Packers were placed in the seven boreholes for a series of sinusoidal pumping tests, which were conducted at the upper and the lower zone of two boreholes, i.e., an oscillatory hydraulic tomographic survey (Cardiff et al., 2013). Lavenue and de Marsily (2001) then employed the pilot point inverse method to characterize the K-field in the Culebra dolomite formation, using these data sets and available geologic facies data. Their characterization was limited to the horizontal variability of the formation.

Vesselinov et al. (2001a, 2001b) conducted pneumatic tomography in unsaturated fractured tuffs at the Apache Leap Research Site (ALRS) in three-dimensions using three cross-hole pneumatic injection tests performed by Illman and Neuman (2001, 2003). They used the pressure records from these tests to estimate equivalent permeability and porosity values, as well as their heterogeneous distributions. The results of the pneumatic tomography were compared to kriged permeability fields based on single-hole pneumatic injection tests (Chen et al., 2000) and were found to share a similar internal structure. In addition, air permeability estimates obtained through pneumatic tomography were compared to single-hole air permeability estimates along several boreholes, yielding a general correspondence between the two estimates. Nevertheless, they used the pilot point approach for inverse modeling of the borehole tests. The results were not optimal.

In a different study, Bohling et al. (2007) assessed steady shape hydraulic tomography (Bohling et al., 2002) in an alluvial aquifer at the Geohydrologic Experimental and Monitoring Site (GEMS) of the Kansas Geological Survey. Steady shape refers to the period during a pumping test where the hydraulic gradients have reached steady state, but the hydraulic heads have not (Bohling et al., 2007). They analyzed a total of 23 pumping tests performed at discrete intervals within two wells that were several meters apart. Between these two wells were two monitoring wells with six vertical observation points that are all aligned in a 2-D plane, to record the drawdown responses. The tomographic analysis produced 1-D vertical profiles of K between the two pumping wells, which agreed reasonably with profiles obtained from a forced-gradient tracer test and direct-push permeameter tests. Bohling and Butler (2010) followed up on their own study showing that heterogeneity in hydraulic parameters could not be obtained perpendicular to the planar configuration of wells in which they conducted the hydraulic tomography survey.

That is, based on their inverse approach, Bohling and Butler (2010) reported that heterogeneity in hydraulic parameters could not be estimated accurately in areas where drawdown data or other information is not available. They employed the pilot point approach and zonation approach to performing model inversion. The resulting shortcomings of such models for HT analysis have been well documented in Huang et al., (2011) and Zha et al., (2017). As a result, analyses of their examples yielded undesirable outcomes.

The power of HT analysis has been recognized after Yeh and Liu (2000) formally introduced a hydraulic tomography technology that allows the use the sequential pumping tests data to image fully three-dimensional heterogeneity in a synthetic aquifer vividly. Similar to the iterative geostatistical technique developed by Yeh et al. (1996) to successively linearize the nonlinear relationship between hydraulic pressure head and parameters, such as hydraulic conductivity (K), they developed the Sequential Successive Linear Estimator (SSLE) for three-dimensional steady-state hydraulic tomography (SSHT) analysis, which jointly inverts multiple pumping tests to map the K-field and corresponding uncertainties. They proved that processing of the data sets from the tomographic survey tests through a consistent mathematical model could yield detailed spatial distribution of hydraulic properties of the aquifer, patterns of connectivity of highly conductive zones, locations of low conductive zones and the uncertainties associated with the spatial distribution.

Then, Zhu and Yeh (2005) extended the SSLE for transient analysis. Their work showed promising results on utilizing transient HT (THT) to characterize accurate estimates of both K- and specific storage (S_s)-fields (or "tomograms" from now on). Since then, geostatistics-based inverse methods have been extensively used for HT data interpretation by several research groups (e.g., Li et al., 2005, 2007; Illman et al., 2007, 2010; Castagna and Bellin, 2009; Liu et al., 2014; Berg and Illman, 2011a, 2011b; Cardiff et al., 2012; Schöniger et al., 2012; Lee and Kitanidis, 2014).

A field example of the use of a geostatistically-based inverse model for hydraulic tomography was first published by Straface et al. (2007), who analyzed six pumping tests performed sequentially within a six-well network, using the transient hydraulic tomography (THT) code developed by Zhu and Yeh (2005) to estimate the heterogeneous transmissivity (T) and storage coefficient (S) tomograms in two-dimensions. Despite the small number of wells, they concluded that the T and S tomograms were reasonable representations of the aquifer based on the geological setting of the site. However, no attempts were made to validate the tomograms.

At the Krauthausen test site in Germany, Li et al. (2008) employed a geostatistical inverse approach to jointly analyze steady-state drawdown and borehole flowmeter data from multiple pumping tests to estimate the K distribution in three-dimensions. They found that jointly inverting both steady-state drawdown and flowmeter data produced an improved 3-D structure when compared to just inverting pumping test data.

The first steady state hydraulic tomography in unconfined aquifers was performed by Cardiff et al. (2009) using nine pumping tests at the Boise Hydrogeophysical Research Site (BHRS) to estimate the distribution of depth-averaged K. They found that the K tomogram contained expected geological features. In addition, the uncertainty bounds on the estimation indicated that K was well constrained within the central portion of the research site, where the pumping and observation well network was located.

Mao et al. (2011) advocated that the classical analysis of unconfined aquifer tests based on gravity delayed water table response theory (Neuman, 1975) is inappropriate (Yeh et al., 2012). The classical theory assumes instantaneous release of water from the falling of the water table. It attributes the S shape of observed drawdown to transition from horizontal to vertical and to horizontal flow. To the contrary, Mao et al. (2011) argued that transition of water release mechanics, from aquifer elastic effects to slow drainage of water from unsaturated zone, and from falling of the water table, is the key to the S shape drawdown. Consequently, Mao et al. (2013a) developed HT for unconfined aquifers.

A large-scale application of THT in fractured rock was demonstrated by Illman et al. (2009) at the Mizunami Underground Research site in Japan. Using two cross-hole pumping tests, they estimated the 3-D distribution of K and S_s as well as their uncertainties. This was the first application of 3-D hydraulic tomography in the field which utilized transient drawdown data. Several continuous high K and low S_s zones were identified and interpreted as possible fault zones. This field investigation highlighted the potential use of hydraulic tomography in fractured rock environments to identify hydraulic connections between boreholes. However, the evaluation of the K and S_s tomograms was limited to available fault data, several drawdown data from one of the cross-hole pumping tests, and coseismic responses in wells. That is, the validity of K and S_s tomograms was not confirmed rigorously through the prediction of independent drawdown inducing events, such as shown previously by Illman et al. (2007, 2008, 2010), Liu et al. (2007), and Berg and Illman (2011a) through sandbox studies. After the work by Illman et al. (2009) on the fractured granite field site, Zha et al., (2015; 2016) included two more pumping test data sets from both sides of the geologically mapped low permeability fault zone into the SLE analysis. They were able to map the detailed irregular shape of the fault zone and found there are local scale high permeability zones in this large-scale fault zone. Zha et al. (2017) also demonstrated that the estimated K and S_s distribution in this fractured granite site could lead to a satisfactory prediction of flow field of an independent pumping test. While these studies confirm the usefulness of HT for mapping fractures and faults in granite rocks on a scale of kilometers, Sharmeen et al. (2012) proved that HT could be used to image microcracks in laboratory rocks. These studies call into question the popular notion (Konikow and Bredehoeft, 1992) that groundwater models cannot be proven or validated.

To further test and validate HT capability, Berg and Illman (2011b) performed three-dimensional hydraulic tomography at the NCRS using short-term sequential pumping tests. They found that HT yielded better results when compared to the inverse modeling of individual pumping tests. The results suggested that the HT results might be further improved by extending the duration of the sequential pumping tests, especially in delineating regions with low hydraulic conductivity. Berg and Illman (2013) also showed that steady state hydraulic tomography is also a viable approach at the NCRS.

More recent studies have shown that when pumping test data are scarce, the geostatistical inversion approach yields overly smooth parameter fields (e.g., Cardiff et al., 2013; Illman et al., 2015). In particular, in the field studies by Cardiff et al. (2013) and Berg and Illman (2011a, 2013, 2015), the geostatistical model yielded K estimates that are inconsistent with geological knowledge for the areas where no pumping and observation data are available. Cardiff and Barrash (2011), through a synthetic study, and Berg and Illman (2015), through a field investigation, have tried to estimate the K tomograms conditioning on prior information of aquifer heterogeneity, such as permeameter K data, in order to improve the consistency of K estimates with geological knowledge.

However, the improvements of K tomogram depend on the availability of hard data and could potentially cause the prediction performance to deteriorate if the local-scale or other data used to improve the estimated parameter fields contain errors (Berg and Illman, 2015).

While it is possible to collect more data for inverse analysis, more efforts are required to obtain additional hydraulic response data or other complementary information (e.g., flux measurements, geological, concentration, temperature and geophysical data) other than pressure heads to calibrate an inverse model. For example, Li et al. (2008) jointly inverted the steady state depth-averaged drawdown HT data and the vertical profile of relative K data obtained from flowmeter tests from fully-screened wells. Brauchler et al. (2012) assessed a sequential inversion approach based on hydraulic and seismic tomography at a field site in Germany. Zha et al. (2014) developed a new approach that can incorporate flux measurements in HT analysis and demonstrated significant improvements to K estimates through a two-dimensional synthetic study. Through a cross-correlation analysis, Tso et al. (2016) showed the benefits of utilizing flux measurements, in addition to drawdown data, in HT surveys through a three-dimensional synthetic case. Soueid Ahmed et al. (2014) and Zhou et al. (2014) conducted synthetic studies to jointly interpret self-potential and pressure head data for K estimation and illustrating the value of self-potential data. Yet in another synthetic study, Soueid Ahmed et al. (2015) presented an image-guided inversion approach to incorporate geological structure information into SSHT analysis. It uses a weighted matrix that contains structure information to regularize the inversion of geophysical or pressure head data.

Geological data are commonly available from outcrops, borehole logs or core samples extracted through drilling. Practically, the geological layer structures do not necessarily represent the zoning of hydrogeological properties (Meyer et al., 2014), due to intralayer heterogeneity and providing no direct hydraulic information (Carrera et al., 2005). Still, geological models are convenient to provide insight into the geological variability the ground and to conceptualize the groundwater flow systems (Koltermann and Gorelick, 1996; Martin and Frind, 1998; Refsgaard et al., 2012). To investigate the utility of geological models in HT, Illman et al. (2015) closely compared the performance SSHT, based on the geostatistical inversion approach, to those from the geological zonation model with perfectly known stratigraphy using the same amount of data. One key finding from the work of Illman et al. (2015) was that when the geological model is perfect, it can yield calibration and validation performances that are comparable to the highly parameterized geostatistical model. In parallel, Schöniger et al. (2015) examined the issue of groundwater model complexity and experimental effort through a Bayesian model selection analysis. Schöniger et al. (2015)'s results indicated that aquifer characterization via HT does not necessarily require an inverse approach based on geostatistics. Instead, an approach based on geological zonation may be more robust, but only if the zonation is geologically accurate.

An important assumption in the works of Illman et al. (2015) and Schöniger et al. (2015) was the perfect knowledge of zonation models based on the geological information. However, such information is impossible to obtain with currently available technology. Therefore, to investigate the issue of utilizing inaccurate geological models for HT analysis and using them as prior information in geostatistics-based HT approach, Zhao et al. (2016) conducted a model comparison study involving four geological models of different accuracies. It showed mixed results in terms of model calibration and validation. Results show that geological models built based on the accurate knowledge of stratigraphy from borehole logs or with errors in stratigraphy

could all be well-calibrated due to the compensational effect of estimated parameters for model structure error (Refsgaard et al., 2012), while the K estimates for each unit can be quite inconsistent from the permeameter K measurements, and model validation results are poorer for those geological models with inaccurate stratigraphy information. Moreover, they found that the performance gap between the geological model and geostatistical approaches decreases in terms of model calibration and validation when the number of pumping tests and monitoring locations is reduced. They concluded that using a geological model as prior information in geostatistical inverse models results in the preservation of geological features, especially in areas where drawdown data are not available. Following up on the study by Zhao et al. (2016), Tso et al. (2016) found that using distributed prior mean K models reflecting layer characteristics for geostatistical inversions leads to better K estimates than inversion cases where homogeneous models are used. Zhao and Illman (2017) clearly demonstrated that prior information at locations outside the well field could enhance the estimates of hydraulic properties and predictions of flow, even within the well field at a field site.

Another approach of HT has also been proposed by Brauchler et al. (2003) based on the asymptotic estimation method developed by Vasco et al. (2000), which uses the travel times of the pressure pulses between two boreholes to estimate the distribution of diffusivity, instead of solving the groundwater flow equation directly with given pressure head data to obtain the K and S_s distributions. This travel-time-based HT approach is found to be computationally efficient, and the reconstructed diffusivity tomograms are found to be useful in providing valuable structural information of K distributions through numerous studies (e.g., Brauchler et al., 2003, 2007, 2011, 2013; Hu et al., 2011).

Lastly, as advocated by Yeh et al. (2015), even in case zonation is known perfectly, if the hydraulic characteristics of each zone are unknown and the number of observation wells is limited, the zonal hydraulic properties estimated by conventional methods could be erroneous and the prediction could be biased. A joint inversion of HT, geological, geophysical, and other related information can lead to better results and is anticipated to be the future direction of subsurface characterization.

2.3 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

Compared to the interpretation of borehole cores or cuttings samples, HT is non-invasive and more cost-effective (especially for deep formations and where direct push approaches have difficulty in high-resolution characterization) for delineating heterogeneous parameter values at all locations. Unlike geophysical tomography, HT directly provides an estimation of K- and S_s-values. In addition, it calculates the uncertainty associated with the estimated K- and S_s-values. Prior research has shown that HT data inherently contain more information than single-well pumping tests, and the joint interpretation method is superior to conventional pumping test data analysis methods in delineating the heterogeneities.

A key advantage of the HT technology is the ability to use existing information and infrastructure to reduce costs and reduce uncertainty associated with any site remediation action. For example, at sites with existing pump-and-treat system, historical operational records and water level monitoring data can readily be used in HT analysis. Other available information from past site investigation, such as well logs, geophysical survey data, flowmeter profiles, and flux measurements, can also be used to enhance the accuracy and to reduce the uncertainty of the HT results.

Besides, additional HT data collection may not be necessary with respect to the site characterization objectives. If additional HT data is needed, the results from HT analysis using existing data can be used to optimize the data collection efforts and costs. The final results will be consistent with the existing information utilized.

The HT technology delineates the spatial distribution of K- and S_s-fields, allowing identification of the high-K/aquifer and low-K/aquitard zones at a site. Preferential chemical transport pathways (i.e., high K zones) and potential back-diffusion source zones (i.e., low K zones) can be identified so that targeted remedial actions appropriate for a site can be developed and remediation design thus can be optimized to enhance its performance.

In addition, HT estimates the uncertainty of the delineated K- and S_s-fields. Such information can be used to evaluate the reliability of remedial action and to maximize the reliability of remediation design.

A limitation of HT is that the resolution of results is dictated by the density of pumping wells and observation ports in wells. For example, Yeh and Liu (2000) suggested that spacing of the observation ports in observation wells should be about the average thickness of the heterogeneity to be mapped in the vertical direction. Likewise, the spacing in the horizontal direction should be approximately the horizontal extent of the stratification. They also suggested that pumping at four different locations (depths and directions) would be sufficient enough (i.e., the return of extensive pumping diminishes rapidly, although it is still useful).

3.0 PERFORMANCE OBJECTIVES

The performance of HT in comparison with other conventional site characterization techniques is evaluated using different quantitative and qualitative criteria. These evaluation metrics are summarized in Table 3-1 and discussed in the following subsections.

Table 3-1. Performance Objectives

Performance Objective	Data Requirements	Success Criteria
Quantitative Performance Objectives		
Determine accuracy of HT against conventional site characterization techniques	Measured drawdown from confirmatory pumping tests; Simulated drawdown by models using hydraulic conductivity (K-) and specific storage (S_s-) fields estimated by HT and conventional methods.	• Bias (HT) < bias (conventional methods) • Mean square error of drawdown (HT) < Mean square error of drawdown (conventional methods) • Observed drawdown within one standard deviation of simulated drawdown based on uncertainty of K and S_s from HT
Determine uncertainty of HT against conventional site characterization techniques	Variance of K and S_s estimated by HT and conventional methods	Variance (HT) < variance (conventional methods)
Qualitative Performance Objectives		
Determine consistency of HT results with geologic/lithologic data	K- and S_s-fields estimated by HT; lithologic/geologic data	Spatial distribution of K and S_s from HT is superior to interpretation from geologic/lithologic data at pumping and observation wells
Determine cost-effectiveness of HT against conventional techniques	Cost for implementing HT and conventional techniques	HT is more cost-effective than conventional techniques
Determine the ease of use for HT against conventional techniques	Level of expertise needed to implement HT and conventional techniques	HT does not require higher level of expertise for implementation in comparison to conventional techniques
Determine the capability of identifying potential low permeability zones	Inferred low-permeability zones from HT and conventional techniques	HT did not miss the low-permeability zones inferred from conventional techniques using data from pumping and observation wells

Notice that the true K and S_s distribution of a site are virtually unknown. We can only qualitatively compare general trends of the K- and S_s-fields derived from HT and classical approaches along boreholes against geological or geophysical well logs. Since the ultimate goal of aquifer characterization is to facilitate better prediction of aquifer responses, our quantitative performance evaluations focus on the prediction performance of these methods.

3.1 PERFORMANCE OBJECTIVE: DEMONSTRATE HIGHER ACCURACY OF HT AGAINST CONVENTIONAL SITE CHARACTERIZATION TECHNIQUES

The accuracy of a joint hydraulic conductivity (K)-specific storage (S_s)-field is determined by how well a model with such spatial distributions of K- and S_s-values can predict the hydraulic response observed during a pumping test when compared with the pumping test on which the K-S_s-field was estimated. At each demonstration site, several pumping tests were conducted.

Results from a set of pumping tests were used to generate the K-S_s tomograms. Data from the remaining pumping tests (confirmatory tests, which are not used in the HT analysis) were used to evaluate the accuracy of the K-S_s-fields estimated by HT and conventional techniques. The accuracy of HT and other conventional methods were compared quantitatively using the following three metrics:

1. Bias in predicted drawdown versus observed drawdown
2. Root-mean-square (RMS) of drawdown prediction errors
3. Drawdown prediction errors relative to the prediction uncertainty due to K-S_s-field estimation uncertainty

3.1.1 Data Requirements

The drawdown observed during confirmatory pumping tests was compared with the drawdown simulated by modeling the K-S_s-fields estimated by HT and other conventional methods. The bias and the RMS of the prediction errors were the data used to assess the accuracy of the HT relative to other conventional methods.

In addition, HT provides uncertainty statistics for the estimated K-S_s-field. The first-order second-moment method was applied to the K-S_s-field to evaluate the drawdown prediction uncertainty. The observed drawdown was compared with the Monte Carlo simulations of drawdown to assess whether the prediction error was within the limit of prediction uncertainty.

3.1.2 Success Criteria

The objective is met if (1) the bias and RMS of the errors in the predicted drawdown using the HT estimated K-S_s-field are smaller than the predicted drawdown bias based on conventional methods and (2) the drawdown prediction error is within the prediction uncertainty resulting from the uncertainty of the K-S_s-field from HT.

3.2 PERFORMANCE OBJECTIVE: DEMONSTRATE LOWER UNCERTAINTY OF HT AGAINST CONVENTIONAL SITE CHARACTERIZATION TECHNIQUES

Estimation variance of K- and S_s-values is a measure of the uncertainty associated with estimation methods. Comparing the estimation variances of various methods allows an assessment of the reliability of the parameter estimation methods.

3.2.1 Data Requirements

The estimation variance of K and S_s computed by HT, and other conventional parameter estimation techniques were compared. The conventional parameter estimation technique to be used for comparison is the first-order approximation method used in the commonly used parameter estimation software, PEST. The data to be considered include the confidence limits determined by PEST.

3.2.2 Success Criteria

The objective is met if the estimation variance associated with HT is smaller than that associated with other conventional methods.

3.3 PERFORMANCE OBJECTIVE: ILLUSTRATE CONSISTENCY OF HT RESULTS WITH LITHOLOGIC/GEOLOGIC DATA

HT calculates the K- and S_s-values at the demonstration sites. Comparing the regions of high K values with regions where relatively coarse-grained materials are located provides a qualitative assessment of the reasonableness of the HT results. Similarly, relating regions with low K values to regions of relatively fine-grained materials provides another qualitative assessment of the HT performance.

3.3.1 Data Requirements

The K- and S_s-values in various regions computed by HT were compared with the available geologic and lithologic data.

3.3.2 Success Criteria

The objective is met if (1) the regions with high K values are consistent with the regions with relatively coarse-grained materials and (2) the regions with low K values are consistent with the regions with relatively fine-grained materials.

3.4 PERFORMANCE OBJECTIVE: ILLUSTRATE COST-EFFECTIVENESS OF HT AGAINST CONVENTIONAL TECHNIQUES

Comparing the costs of HT with conventional site characterization methods in regard to the quality of resulting site characterization provides a qualitative assessment of the cost-effectiveness of the HT results.

3.4.1 Data Requirements

The itemized and total costs of HT and conventional site characterization methods are considered. The K-S_s parameter distribution delineated by HT and conventional methods will be used.

3.4.2 Success Criteria

The objective will be met if HT is more cost-effective than conventional methods.

3.5 PERFORMANCE OBJECTIVE: ILLUSTRATE THAT HT IS 'USER-FRIENDLY.'

The objective is to illustrate that HT is 'user-friendly' and can be readily applied to other sites.

3.5.1 Data Requirements

The ease-of-use information for HT and conventional site characterization methods will be used.

3.5.2 Success Criteria

The objective will be met if HT is 'user-friendly' and can be readily applied to other sites.

3.6 PERFORMANCE OBJECTIVE: ILLUSTRATE THAT HT IS ABLE TO IDENTIFY LOW-CONDUCTIVITY ZONES

The objective is to illustrate that HT is able to identify areas with low-conductivity zones that might not have been identified using conventional techniques.

3.6.1 Data Requirements

Regions with low hydraulic conductivity values delineated by HT and conventional methods will be used.

3.6.2 Success Criteria

HT is able to identify low-conductivity zones that have been identified using conventional methods. In addition, HT might be able to identify low-conductivity zones that have not be identified by conventional methods.

4.0 SITE DESCRIPTION

The HT technology was demonstrated at the North Campus Research Site (NCRS) and the U.S. Air Force Plant 44 (AFP44) site. These sites, their history, and relevant hydrogeologic information are described in the following sections.

4.1 SITE LOCATION AND HISTORY

4.1.1 NCRS at UW, Canada

The NCRS is located on the UW campus in Waterloo, which is approximately 100 km west of Toronto, Ontario, Canada (Figure 4-1). The site has been historically utilized as an on-campus field site for Earth 671 (Field Methods in Hydrogeology) and Earth 458 (Physical Hydrogeology) for the hydrogeology graduate students from the Department of Earth & Environmental Sciences at the UW.

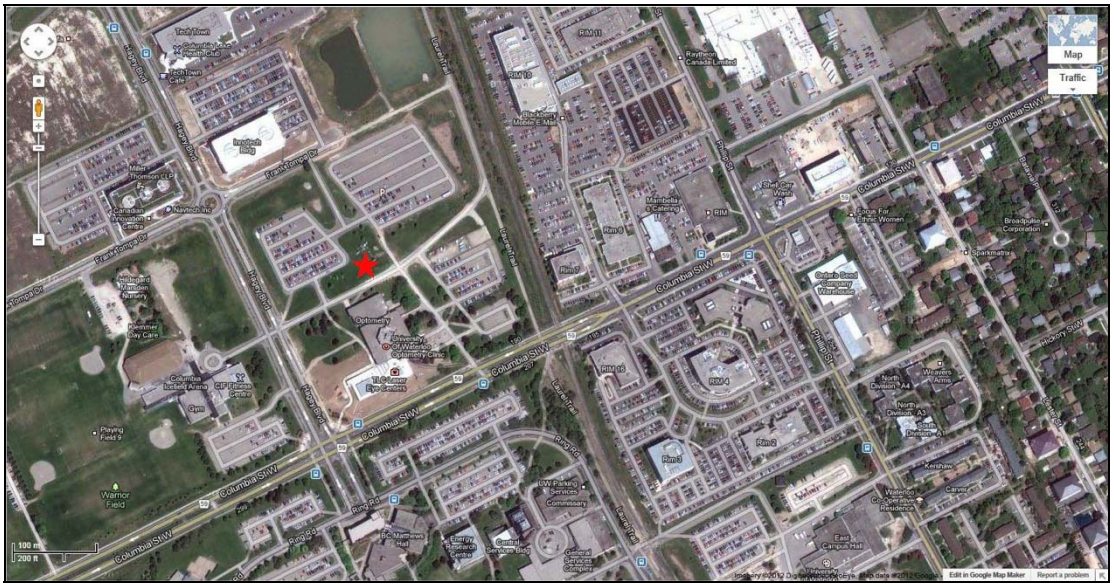


Figure 4-1. Map of the Northern Part of the University of Waterloo Campus

The location of the NCRS is indicated with a red solid star (modified image from Google Maps).

4.1.2 AFP44 in Tucson, AZ

The AFP44 site is a recalcitrant environmental site located in the southern portion of the Tucson International Airport Area (TIAA) CERCLA site, approximately eight miles south of downtown Tucson, Arizona. Figure 4-2 shows a map of the area indicating the location of the 1266-acre AFP44 site. The TIAA has been included in the National Priority List (a.k.a. Superfund) by the U.S. Environmental Protection Agency and is currently under the jurisdiction of EPA Region IX. The site is located in the Tucson Basin, with an average annual precipitation of 11.59 inches between 1981 and 2010. The regional groundwater flow direction is to the northwest. A brief history of the site and a chronology of events can be found on the Arizona Department of Environmental Quality (ADEQ)'s website.

The AFP44 site has been historically utilized as a federally-owned weapons manufacturing facility operated under contract through Hughes Missile Systems (later acquired by the Raytheon Company) since 1951. The historical industrial processes conducted at the AFP44 site have resulted in two major commingled plumes of Trichloroethylene (TCE) and 1, 4-dioxane contamination in both its groundwater and vadose zone.

Throughout the more than thirty years of site history, numerous site investigations have been conducted by the USGS (e.g., Hanson and Benedict, 1994; Houser et al., 2004; Tillman, 2009) and various consultants (e.g. AECOM, 2010, 2011, 2012; URS, 2013a, 2013b, 2014). There is, however, no unifying geological framework developed for the entire TIAA area, which has caused different groups working in the same area rarely referring to the outcomes from others.

Numerous wells have been installed at the site by various agencies. To obtain a complete list of well locations and screen intervals, we have synthesized and reconciled information from several key data sources (Earth Tech, 2007; HydroGeoLogic, 2012; Montgomery & Associates, 2015). A regional mapping of hydraulic conductivity based on the texture of core logs has been performed (Zhang and Brusseau, 1998), and the resultant field has been used to run groundwater and transport models (Zhang and Brusseau, 1999).

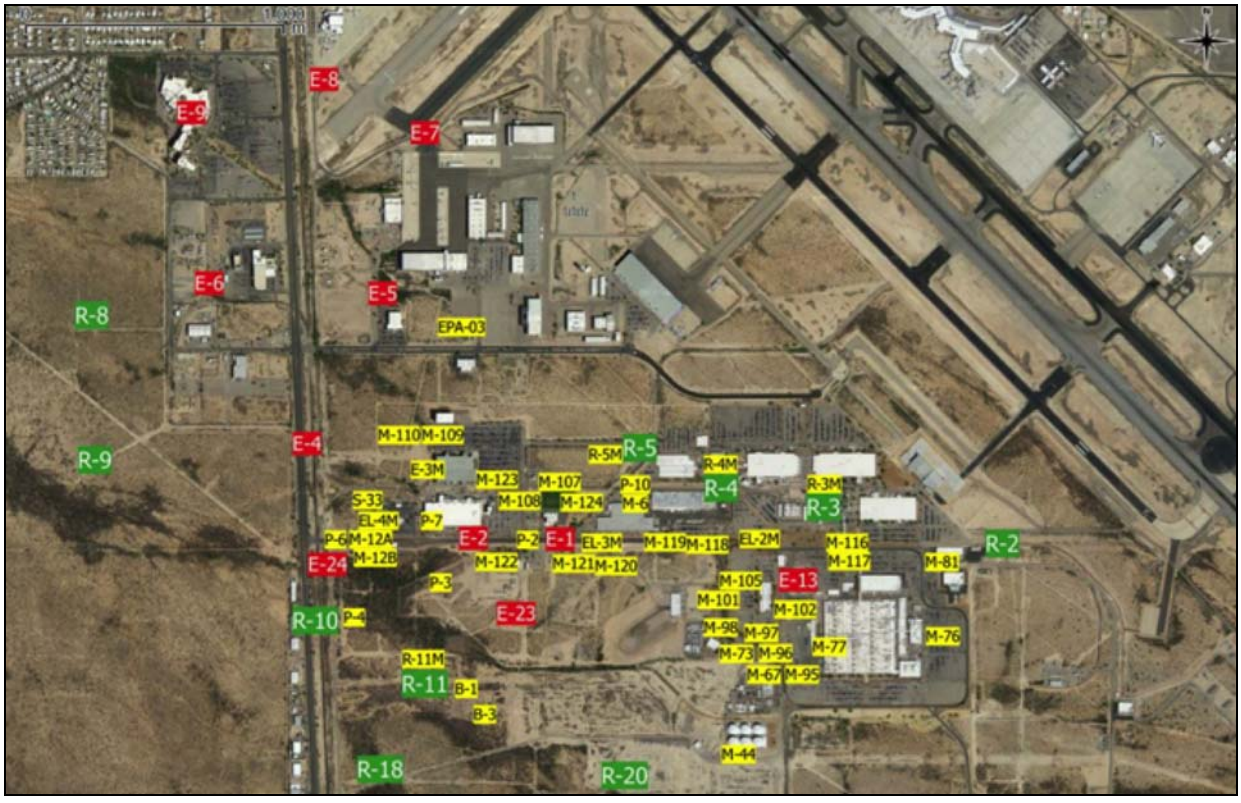


Figure 4-2. Map of AFP44 Site in Tucson, Arizona

4.2 SITE GEOLOGY/HYDROGEOLOGY

4.2.1 NCRS at UW, Canada

The surficial geology of the NCRS area is glacis fluvial in origin and is highly heterogeneous. The field site is located within the Waterloo Moraine, which is an interlobate feature composed of kettle and kame deposits, which contain alternating layers of till and glaciofluvial material (Karrow, 1993). Site-specific geology has been described by Sebol (2000) and Alexander et al. (2011). The main feature of the site is an "aquifer zone" located approximately 8 to 13 m below ground surface (mbgs). This zone consists of two high K units that are separated by a discontinuous low K layer. The upper aquifer is composed of sand to sandy silt, and the lower aquifer is composed of sandy gravel. The low K unit separating the two aquifers is discontinuous and is known to provide a hydraulic connection (Alexander et al., 2011). In addition, despite being interpreted as continuous layers, none of the units extend across the entire study site. Situated above and below the aquifer zone are low K silts and clays. At approximately 18 mbgs is the dense Catfish Creek Till, which acts as a hydraulic barrier (Alexander et al., 2011) and is taken to represent the bottom boundary for this study.

Near the ground surface, the aquifer system is generally confined by a laterally extensive upper aquitard layer. However, this aquitard is known to contain stratigraphic windows in some areas (Martin and Frind, 1998). Based on previous pumping tests performed at the site (Alexander et al., 2011), the aquifer at the NCRS behaves as a confined aquifer. None of the drawdown responses observed during previous pumping tests suggest that the main aquifer zone behaves in an unconfined manner, which might indicate the presence of stratigraphic window(s). Water levels collected in the vicinity of the site indicate that groundwater flow is toward the southeast. Depth to water is relatively shallow.

Figure 4-3 shows the distributions of wells from which geological information were obtained and the locations of the cross-sections presented in Figure 4-4. In total, we used borehole logs from 18 pumping and observation wells consisting of different depths, ranging from six meters to 18 meters below ground surface. Based on the soil types and corresponding depth information, 19 different layers representing seven different material types are defined along all boreholes. The layer information between boreholes at different locations was interpolated using the commercial software Leapfrog Hydro (ARANZ Geo Limited), to construct a three-dimensional geological model with dimensions of 70 m × 70 m × 17 m. Four cross-sections (A-A', B-B', C-C', and D-D', in Figure 4-4) are extracted along different directions among the central nine wells to show the interpolated geological layers, as shown in Figure 4-4. Moreover, the locations of wells and screens are also presented for cross-sections C-C' and D-D'. The model reveals that the units are highly discontinuous, contributing to the strong heterogeneity at the site.

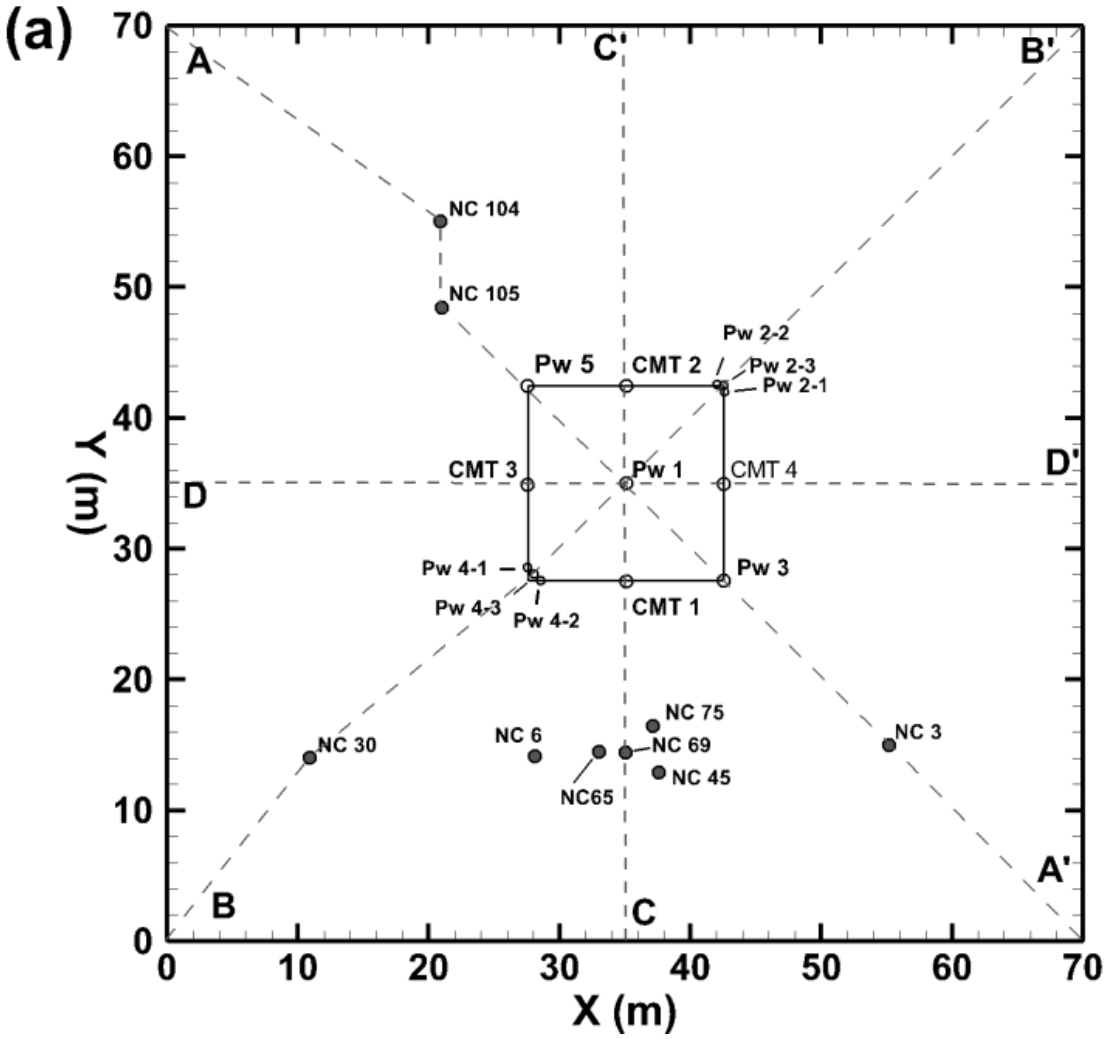


Figure 4-3. Locations of Wells and Cross-sections

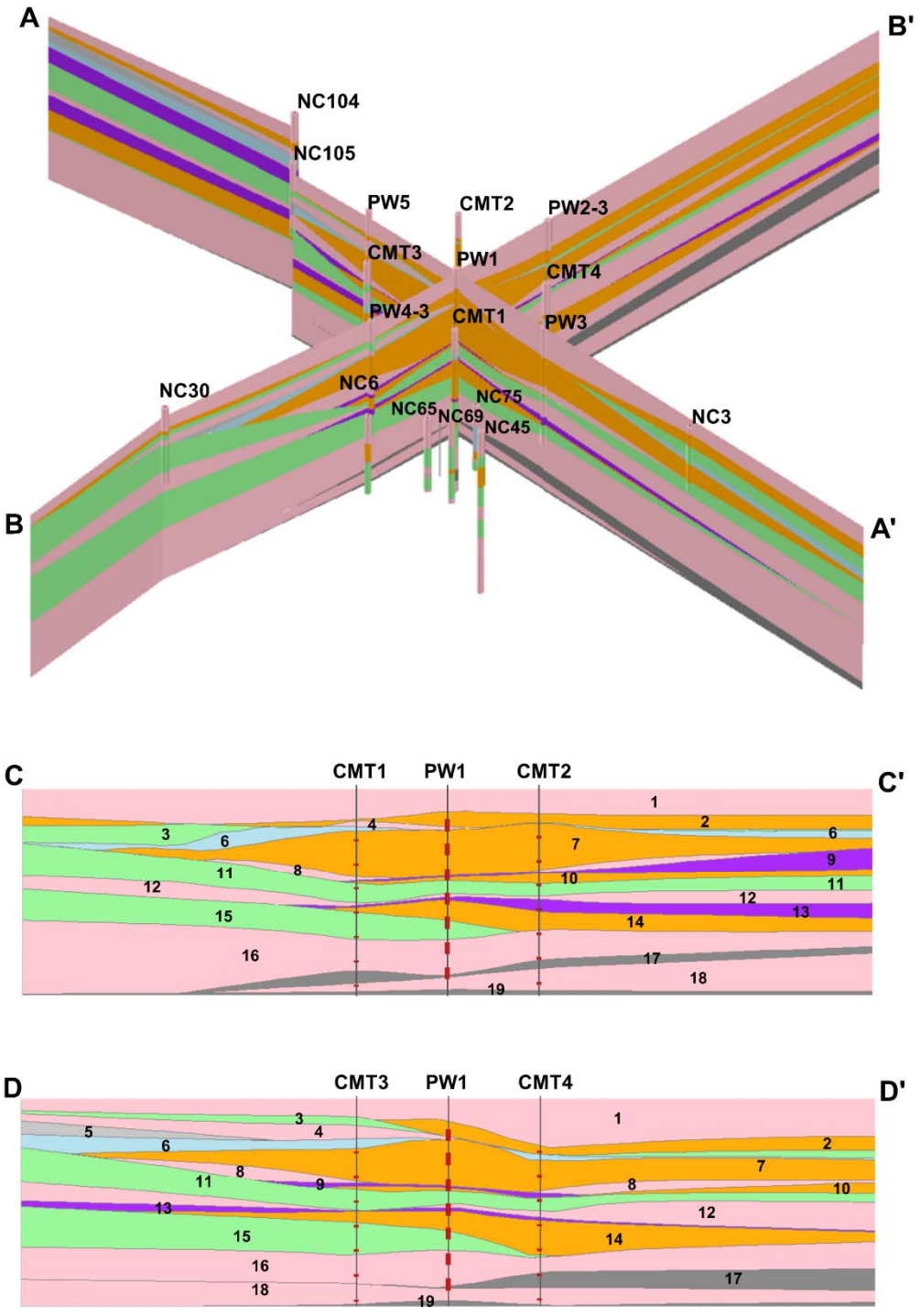


Figure 4-4. Stratigraphic Model of the NCRS

Numbers in cross section C-C' and D-D' indicate the 19 layers of different materials: Clay (1, 4, 8, 12, 16, 18); Silt and Clay (17, 19); Silt (2, 7, 10, 14); Sandy Silt (6, 9, 13); Sand and Silt (5); Sand (3, 11); Sand and Gravel (15). Screened locations are shown on wells depicted in cross sections C-C' and D-D'.

4.2.2 AFP44 in Tucson, AZ

The AFP44 site is situated on the western edge of the Tucson Basin, within the intersection of the large, ancient Cienega Alluvial fan and the Santa Cruz River, both of which are highly heterogeneous systems that have been reworked to result in a complex and unpredictable depositional environment (AECOM, 2012). Groundwater at the AFP44 site is hydraulically controlled by an active remediation system that extracts, treats, and then re-injects the treated water on site. For the past 13 years, the water table at the AFP44 site has risen 80 feet in response to the Pima Mine Road Recharge Project due to the proximity (5 miles) of the AFP44 site to the infiltration ponds.

Figure 4-5 shows a three-dimensional perspective view of the subsurface conditions at AFP44. The study area is underlain by these unconsolidated to semi-consolidated alluvial sediments to at least 600 feet bgs and has been further characterized as belonging to three primary stratigraphic units: the Holocene Alluvium (a few feet to approximately 30 feet bgs), the Fort Lowell Formation (depths down to 220 feet bgs), and the Tinaja beds (below the Fort Lowell Formation to 600 feet) (URS, 2013). Site-specific geology has been described in more detail by AECOM (2012) and URS (2013). There are two aquifer zones identified within these basin-filled sediments, labeled as the semi-confined Upper Zone (UZ) within the Fort Lowell Formation and the confined Lower Zone (LZ), which supplies municipal drinking water to the city of Tucson, in the Tinaja beds (Figure 4-6).

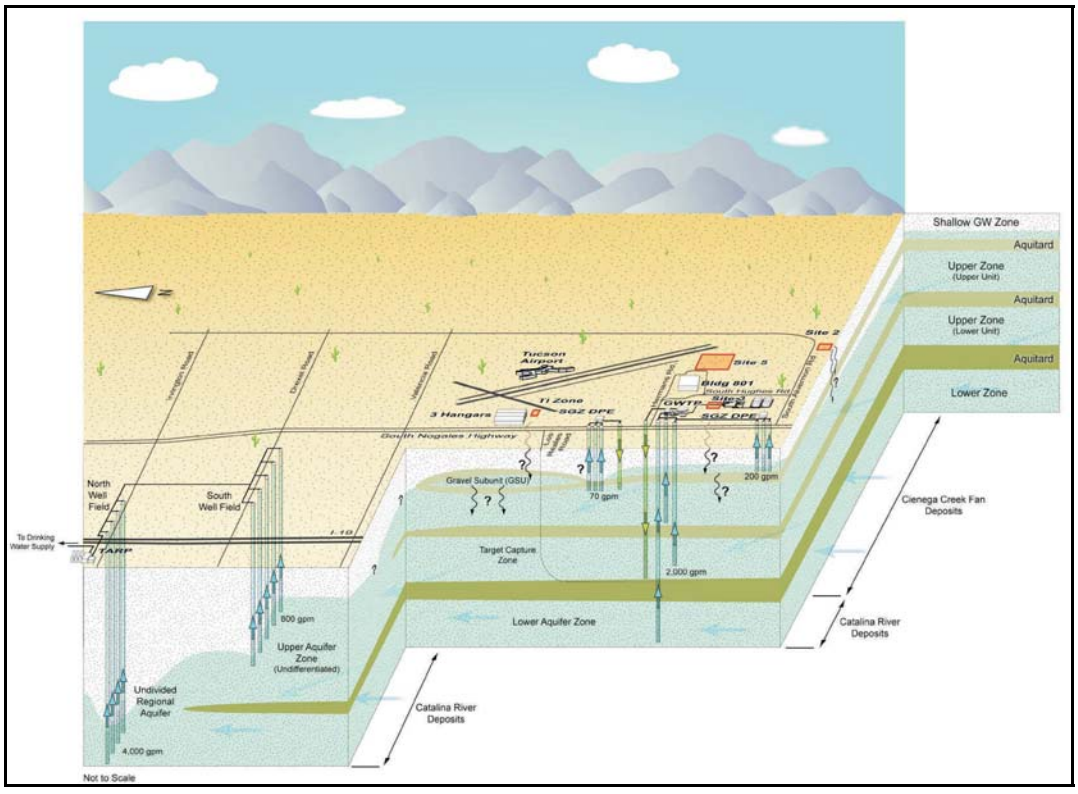


Figure 4-5. Three-dimensional Perspective View of the Subsurface Conditions at AFP44

(source: AECOM)

The UZ is the most productive aquifer unit and, since it contains most contaminants, is the focus of the AFP44 groundwater remediation project. It has an estimated average K of 13 feet/day to 133 feet/day based on data from extraction and recharge wells within the AFP44 site (Earth Tech, 1992). The UZ extends to 220 feet bgs and is further subdivided into two distinct aquifer subunits, the upper zone upper unit (UZUU) ranging from the water table to 160 feet bgs, and the Upper Zone Lower Unit (UZLU) from 160 to 220 feet bgs. The two units are separated by an aquitard that is present over much of the study area that pinches out to the west near recharge wells R-8 and R-9, where it is possibly one undivided UZ aquifer (URS, 2013). Where present, the confining unit is typically 55 feet thick. It hydraulically isolates the UZUU and UZLU aquifers. The UZLU is lithologically similar to UZUU, but it contains a higher percentage of gravel. Most of the extraction wells at AFP44 are screened across both the UZUU and UZLU, producing an average water level within the screen depth interval. However, a larger portion of the pumped water comes from the UZLU. Hydraulic heads are typically lower in the UZLU than the UZUU.

The LZ is separated from the UZ by a confining unit correlated with the Upper Tinaja beds (Leake and Hanson, 1987) that is comprised of a clayey silt and mudstone from the base of the UZ to about 250 to 300 feet bgs. This confining unit pinches out to the west and north of the project area, creating an undivided regional aquifer by eliminating the UZ and LZ separation. The LZ has lower average estimated K of 0.1 to 1.3 feet/day (Earth Tech, 1992), attributed to less coarse-grained sediments and more consolidation and cementation than in the UZ. The vertical hydraulic gradient between the UZ and the LZ is downward.

Therefore, the objective of the HT study is to estimate and delineate the K distribution of UZUU, UZLU, and their separating aquitard at high resolution.

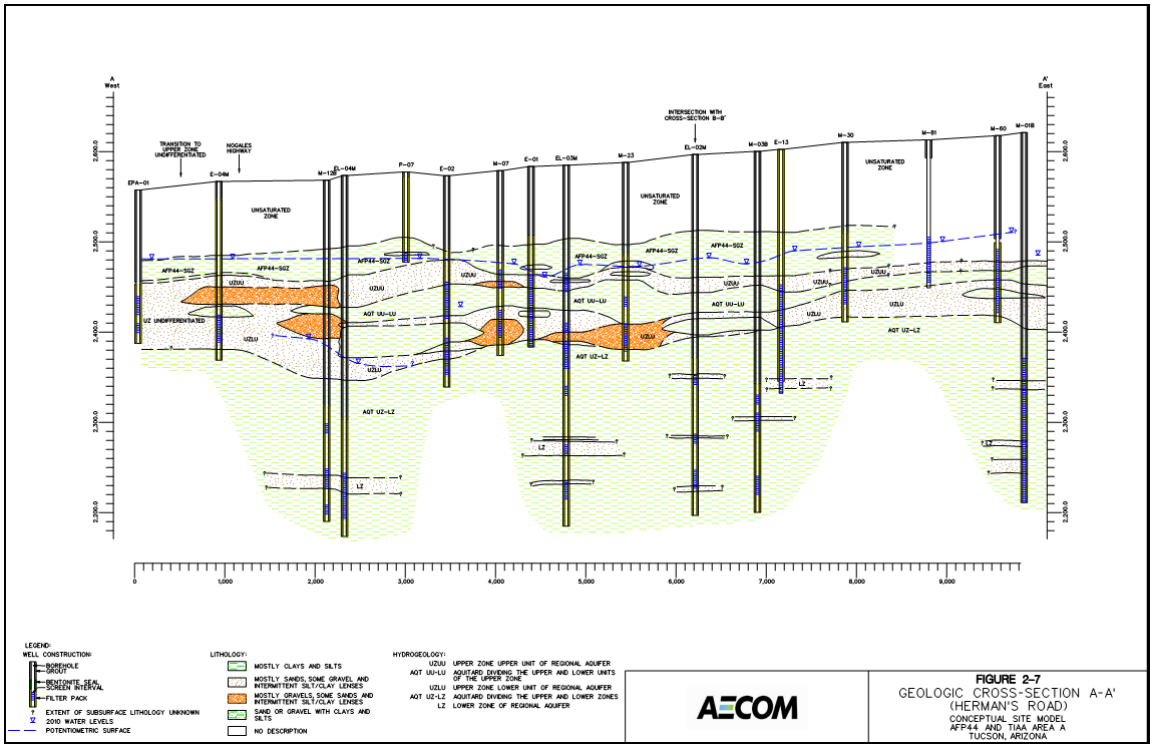


Figure 4-6. Stratigraphic Model of AFP44 Site (from AECOM, 2012)

4.3 CONTAMINANT DISTRIBUTION

4.3.1 NCRS at the UW, Canada

To our knowledge, there are no contaminants at the NCRS except for some salt applied on nearby roads.

4.3.2 AFP44 in Tucson, Arizona

The primary constituents of concern at the site are TCE, 1,1-DCE, 1,4-dioxane, and chromium. These constituents were released to the subsurface via pits, channels, and leaks. The chemicals migrated down through the vadose zone to the groundwater system. They have since migrated into and accumulated in the fine-grained sediments of the aquitard overlying the UZUU, where they continue to serve as a source from back-diffusion. The process of drilling wells through both the UZUU and into the UZLU created a conduit for migration of the constituents of concern down to both subunits of the UZ. Similar migration occurred in the early 1980s between the UZ and the LZ at three production wells. Figure 4-7 shows the lateral extent of the groundwater plume at the AFP44 site.

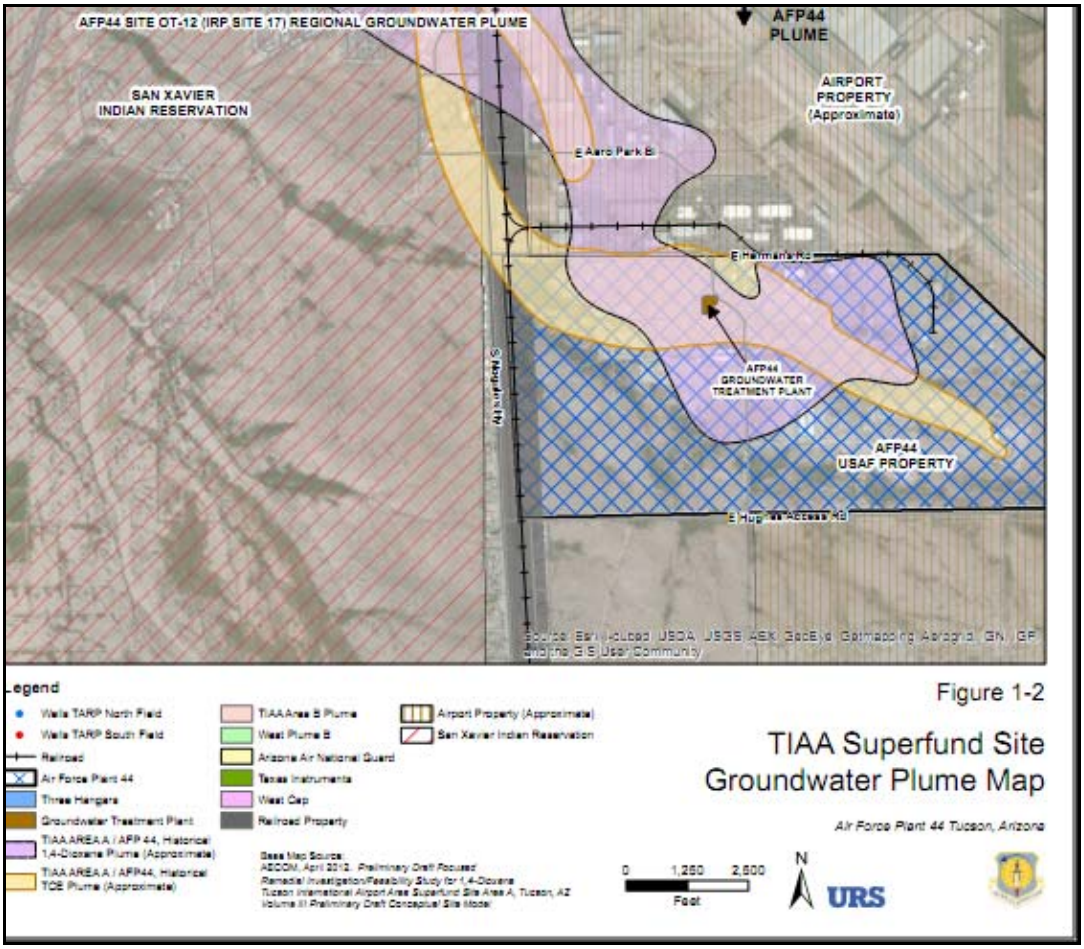


Figure 4-7. Groundwater Plume Map at the AFP Site

5.0 TEST DESIGN

5.1 CONCEPTUAL EXPERIMENTAL DESIGN

The HT pumping tests (HT survey) conducted for this demonstration project were designed to perturb the hydraulic head fields spatially and to measure the corresponding hydraulic responses at multiple locations in response to each perturbation. These tests were performed using the existing well network and site facilities at both the NCRS and AFP 44 sites. At the NCRS, pumping tests were designed to extract (or inject) groundwater from (or to) individual screen intervals in order to generate non-redundant hydraulic stresses at distinct locations. At the AFP44 site, pumping tests were designed to modify the pumping rates at the extraction and injection wells in order to generate non-redundant changes in the spatial distribution of hydraulic stresses. The onsite treatment system requires a minimum flow rate of 2,500 gpm and a maximum flow rate of 5,000 gpm. Therefore, the total extraction rate and total injection rate need to be maintained within this range.

Prior to conducting the HT pumping test at each site, existing site information were reviewed, including stratigraphy data, existing slug and conventional pumping tests, and estimates of K obtained from core samples. Based on this information, initial analytical and numerical models were built to forecast the system response based on existing data. The results were used to select the number of pumping and observation wells, the associated screen intervals, and the pumping test durations for the HT pumping tests. The test program was designed to capture the site heterogeneity in sufficient detail.

Two sets of HT pumping test data were collected at each site. Pressure transducers were installed in the monitoring wells to collect hydraulic response data during the HT pumping tests. These transducers also recorded data before and after the pumping tests to provide information data for removal of background trends in the data. One set was used in the HT analysis to delineate the K and S_s distributions. The resulting K and S_s distributions are used to predict the second set of pumping test data. A comparison of the predicted and observed pumping test responses for the second dataset was used to evaluate the performance of HT. In addition, the delineated K and S_s distributions were compared with existing lithologic information and permeameter test results to evaluate their consistency.

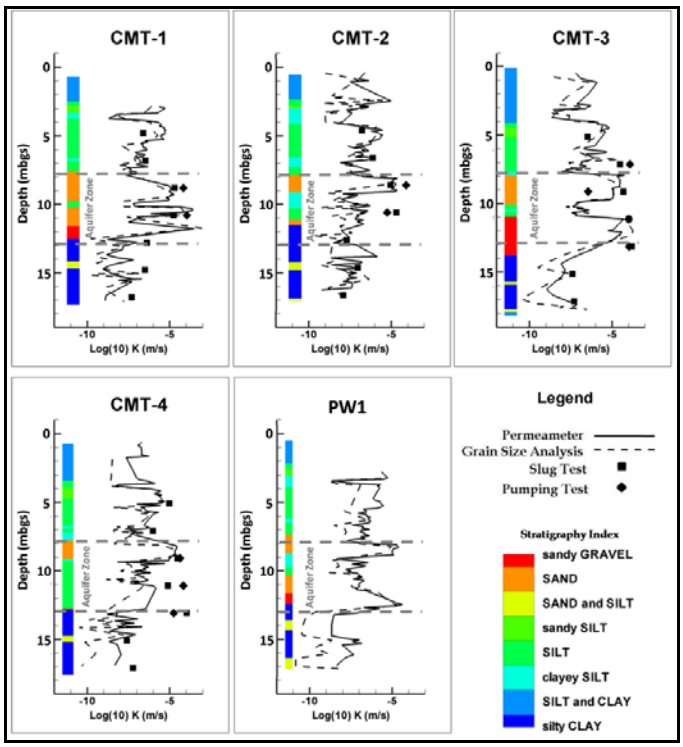
5.2 BASELINE CHARACTERIZATION

5.2.1 NCRS at UW, Canada

The NCRS site has been investigated through pumping tests and other traditional approaches (core sampling, permeameter tests, grain size analysis, slug tests) by Alexander et al. (2011), and Berg and Illman (2011b), as well as through this study. Alexander et al. (2009, 2011) analyzed five continuous soil cores collected during the installation of pumping and monitoring wells. This analysis consisted of detailed core logging, 471 permeameter tests, and 270 grain size tests of core samples. Soil core sample analysis from the previous and current studies shows that main aquifer layers are between seven and 13 meters below the ground surface and that this aquifer zone consists of two high K units separated by a discontinuous low K unit. K-values were estimated from permeameter analyses of core samples and grain size distribution data using empirical relationships.

Several pumping tests along PW1 and slug test data in 28 ports of the four CMT wells have been performed prior to this study. Drawdown data from pumping tests indicate that the permeable unit behaves as a semi-confined aquifer in our study area. Investigations by Alexander et al. (2011) suggested that the low K unit separating the two aquifers is discontinuous and is known to provide hydraulic connections. K values were estimated from the pumping tests and slug test data using analytical solutions based on the assumption of uniform medium. The K estimates from grain size analyses, permeameter analyses, and slug and pumping tests are summarized in Figure 5-1a. Alexander et al. (2011) utilized the permeameter K data to conduct a geostatistical analysis. Figure 5-1b shows the location of core samples that were used for permeameter analysis to create a kriged K-field (Alexander et al., 2011). The geostatistical analysis showed that the site is highly heterogeneous consisting of thin discontinuous beds with abrupt changes in material types. Based on raw permeameter K values, Alexander et al. (2011) estimated the $\sigma^2_{\ln K}$ to be 6.5. The estimated vertical correlation length for the site was approximately 15 cm, and K was found to vary over five orders of magnitude.

a)



The solid and dashed black lines represent laboratory permeameter and grain size estimates, respectively. The square represents slug tests estimates, and the diamond pumping test estimates. The horizontal gray dashed lines indicate the approximate location of the "aquifer zone."

b)

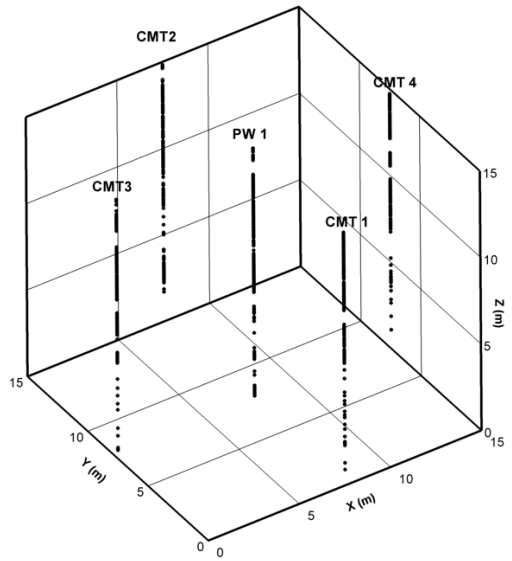


Figure 5-1a,b: a) K estimates along 5 Boreholes at the NCRS (modified after Alexander et al., 2011); b) Location of Core Samples used for Permeameter Analysis to Create the Kriged K-field in Alexander et al. (2011) and Utilized to Condition Some of the Models in this Study

5.2.2 AFP44 in Tucson, Arizona

The available information is from the preliminary site investigation work completed by the AFP44 consultants (AECOM, 2012; Earth Tech, 2007; HydroGeoLogic, 2012; Montgomery & Associates, 2015; URS, 2013a, 2013b, 2014). The project team has reviewed the existing site information (including the stratigraphy data), existing slug and conventional pumping tests, estimates of K obtained from core samples, and an existing groundwater model. The baseline characterization activities include groundwater level measurements at all the extraction and observation wells. Baseline monitoring of groundwater levels at all the wells was conducted prior to the initiation of a HT survey. A network of pressure transducers was installed at the AFP44 site. The pumping rates at extraction wells E-01, E-02, E-03, EL-03, E-04, E-05, E-23, and E-24 were monitored. Selected observation wells were monitored. Available site data also includes a draft numerical groundwater flow and transport model. The model was examined to provide reference information for designing the details of the HT pumping tests.

5.3 DESIGN AND LAYOUT OF TECHNOLOGY COMPONENTS

5.3.1 NCRS at UW, Canada

The NCRS is instrumented with a total of seven wells (PW1, PW3, PW5, CMT1, CMT2, CMT3, and CMT4) and two well nests (PW2 and PW4) in a square pattern (one at each corner, one at the center of each face, and one in the center) measuring 15 m × 15 m. Figure 5-2 is a schematic layout of the wells.

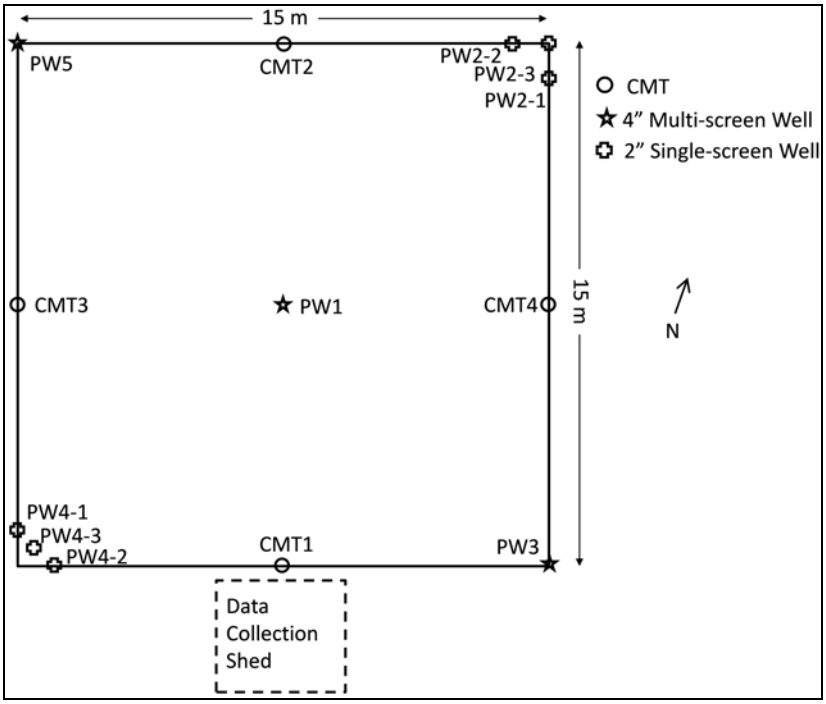


Figure 5-2. Two-dimensional Plan View Showing Well Locations

The circles indicate the location of the 4 CMT wells, the stars the location of the multiscreen wells, and the crosses the location of the 2" well nests.

Continuous multichannel tubing (CMT) wells (Solinst Canada Ltd.), containing seven channels each (seven screened intervals), are used strictly as observation wells and are installed in between the four corners of the square pattern (Figure 5-2). The screened intervals of the CMT wells are spaced 2 m apart with the upper screens located between 4.5 to 5.5 mbgs, and the deepest ports are 16.5 to 17.5 mbgs. The remaining five wells are pumping wells (PW), three of which are multiscreen wells (PW1 contains eight screens, PW3 and PW5 contains five screens). These multiscreen wells are 10 cm in diameter and contain screens of 1 m in length. Each screen is separated from adjacent screens by 1 m of solid PVC pipe. PW1 extends to approximately 18 mbgs; and PW3 and PW5 to 12 mbgs. PW2 and PW4 are well nests consisting of three separate wells each (5 cm diameter). Each well in the nest has one screen that is 1 m long. Screen elevations at the midpoint for PW2 are 4, 7, and 8 mbgs, and screen elevations for PW4 are 5, 8.5, and 11.5 mbgs. Bentonite layers were installed between the sand packs around adjacent well screens to provide hydraulic isolation of individual screen intervals, with the exception of PW2 and PW4. In these two well nests, wells are installed in direct contact with the native formation. Figure 5-3 shows the three-dimensional perspective view of the wells and the various screens, as well as bentonite seal elevations at the NCRS.

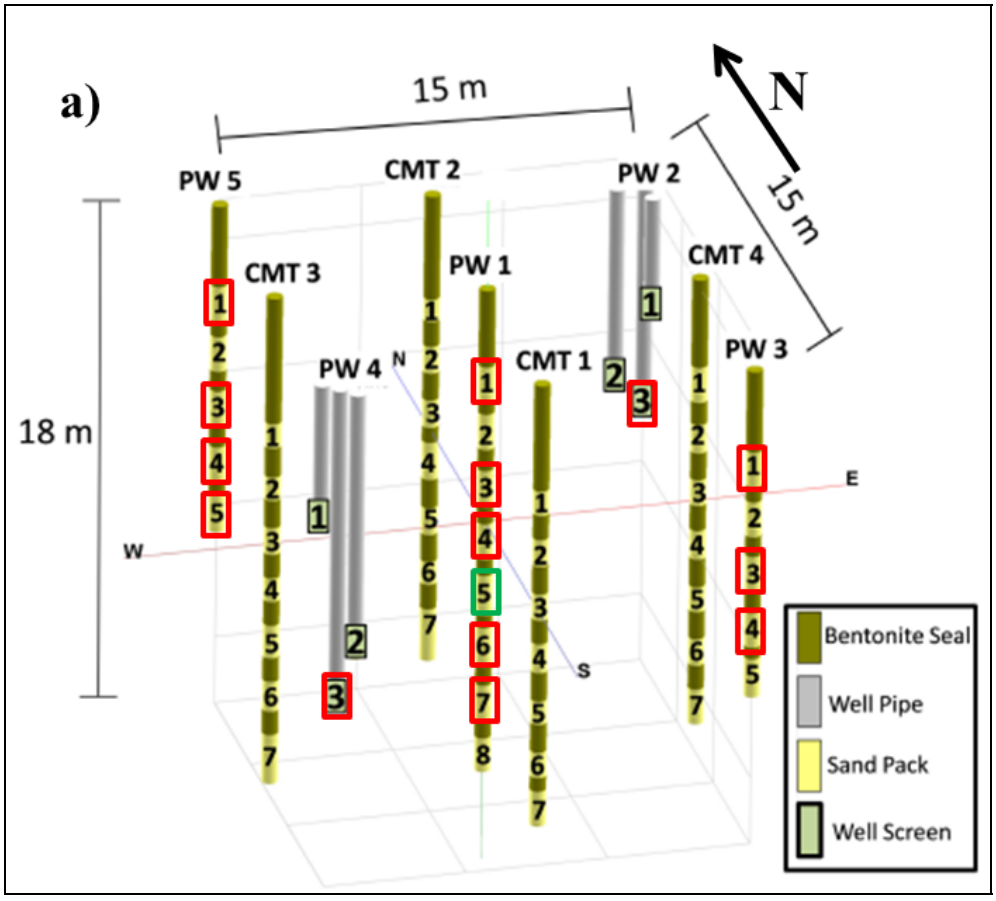


Figure 5-3. Three-dimensional Perspective View of Various Wells and Pumping Locations

A pressure transducer network was installed to collect pressure head data at up to 44 observation ports, depending on the particular pumping test. For CMT well ports, 0-15 psig (model MP100: Micron Systems) pressure transducers were used for monitoring. PW2 and PW4 well nests were monitored with 0-5 or 0-10 psig pressure transducers (model 3001 LT Leveloggers Junior, *Solinst Canada Ltd.*). For the multi-screen wells (PW1, PW3, or PW5), FLUTe (*FLUTe Ltd.*) liners were installed in the well not used for pumping to prevent hydraulic short circuiting between adjacent screens within the well. On the other hand, two liners contained five vented pressure transducers (Level Troll, *In Situ Inc.*) each at locations correspondent to the PW well screens. These liners can be moved between wells due to the similar construction of PW1, PW3, and PW5. When one of the multi-screen wells (PW1, PW3, or PW5) was pumped, the FLUTe liners were installed in the two unpumped wells. When PW2 or PW4 was used for pumping, the FLUTe systems were installed in PW3 and PW5, and a blank FLUTe liner was installed in PW1.

5.3.2 AFP44 Site in Tucson, Arizona

An active groundwater remediation system has been operating at the AFP44 site. The system is comprised of numerous wells for extraction (including dual phase extraction), injection, and monitoring. The focus area for the field-scale demonstration is an approximately 4.6 square-mile area within the AFP44 site. Existing extraction wells, injections wells, and monitoring wells were used in this HT study. A well inventory was initially provided by URS Corporation. During the first site visit in preparation for the first pumping test, we were not able to locate some of the monitoring wells. It was suspected that those wells no longer exist. In addition, some of the monitoring wells have dedicated sampling pumps installed. The inner diameters of the sounding tubes are too small for installing pressure transducers. Some of the monitoring wells were dry. Some wells had too little water remaining and might become dry during pumping tests, so they were not good candidates for monitoring. Meanwhile, new wells were discovered. The locations and names of these new wells were subsequently provided to us. URS informed us that some of the extraction and injection wells are non-operational due to various reasons. There was no plan to repair these wells during the course of this study.

Figure 5-4 shows the extraction wells, injection wells, and monitoring wells available for this study. Figure 5-5 shows the screen intervals of the wells. A network of pressure transducers was installed in selected wells for data collection during the pumping tests. Most of the selected extraction wells and observation wells are completed within the UZ of the regional groundwater zone. These wells are screened at depth intervals between 79 and 230 feet below ground surface (bgs). Several selected extraction wells and observation wells are completed within the LZ. Approximately 18 selected observation wells and one extraction well are screened exclusively in the UZUU. Approximately 15 selected observation wells and six extraction wells are screened across both the UZUU and the UZLU. Approximately four selected observation wells and one extraction well are screened in the LZ. Based on past observations, the hydraulic connection between the different aquifer zones is believed to be minimal due to orders of magnitude differences in the estimated hydraulic conductivities between zones separated by relatively thick confining units.

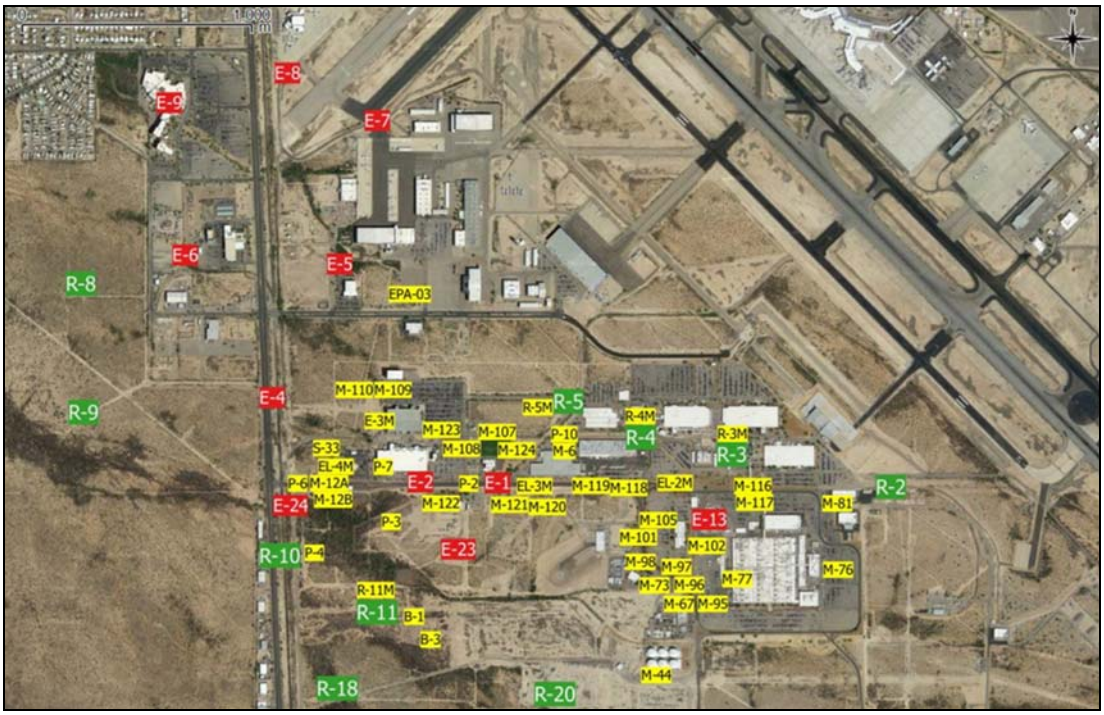


Figure 5-4. Locations of Extraction Wells (red), Injection Wells (green), and Monitoring Wells (yellow)

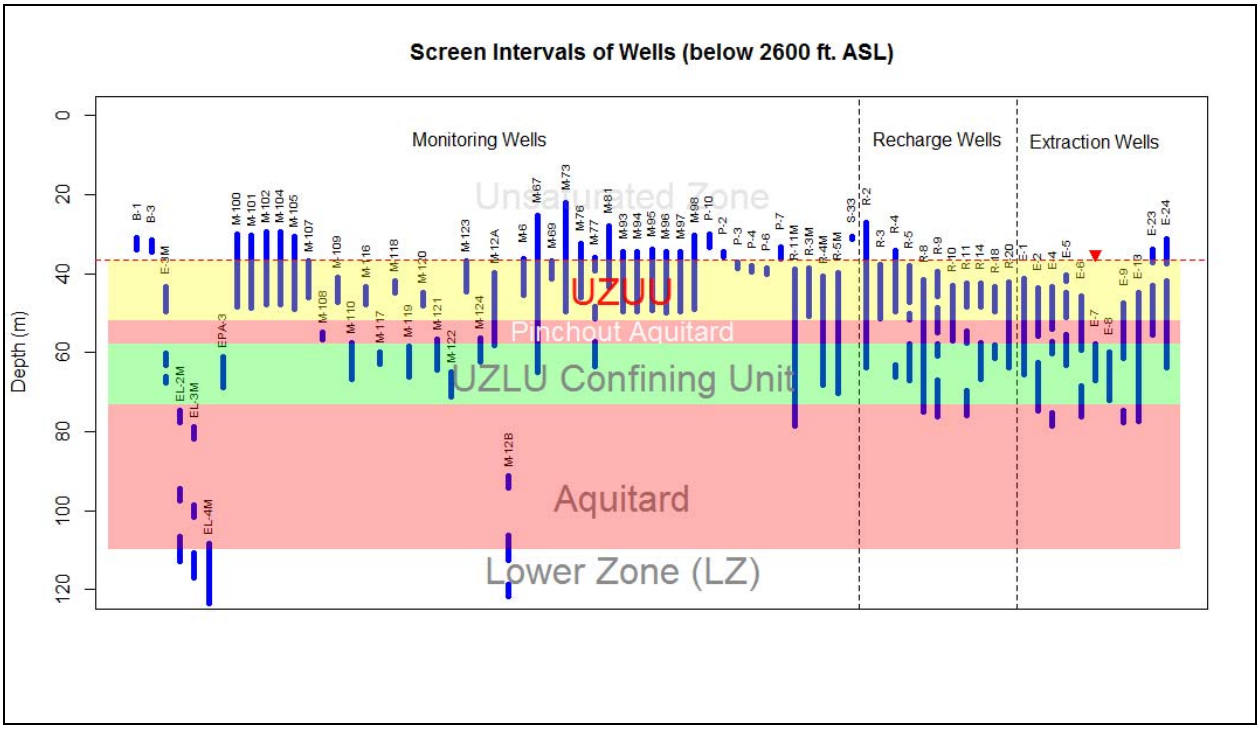


Figure 5-5. Screen Intervals of Wells

5.4 FIELD TESTING

Per instruction from the ESTCP, the field program at the NCRS was conducted first. The experience gained, and lessons learned were used to guide the field program at the AFP44 site.

The work at the NCRS demonstrated that injection of water, as an alternative to groundwater extraction, produces useful data for HT analysis. It also indicated that applying hydraulic stress in low-K zone was important. We also learned that repeating several pumping tests and injection tests were necessary to reduce noise and background influence. As a result, for the demonstration at the AFP44 site, both extraction and injection schemes were perturbed. All wells, including the relatively high-yield and relatively low-yield wells associated with the existing pump-and-treat system, were used. The transducers were installed in the wells for the longest period possible to record the responses of similar hydraulic stress perturbations.

5.4.1 NCRS at UW, Canada

Contractual issues and unfavorable weather conditions were encountered in the spring of 2013. In order to avoid the potential hydraulic influence from the field activities of the University of Waterloo field school program near the NCRS, the initiation of field work at the NCRS was delayed to June of 2013. At the start of the field program, water levels at all screen intervals were measured at all wells. Wells PW2-1 and PW4-1 were dry; pumping tests could not be performed at these two locations.

During the fall of 2013 and 2014, pumping tests were attempted at all non-dry well screen intervals of all pumping wells. For PW1, PW3, and PW5 wells, pumping was performed using a submersible pump (Model SQE05, *Grundfos Canada Inc.*) located between two inflated packers to isolate the target screen. Pumping in PW4 and PW2 well nests was performed using a surface lift pump. Logging of the pressure transducer and barometric data started three days prior to the pumping test to establish baseline hydraulic head levels. Manual water level measurements were taken before the test commenced. During each pumping test, we collected drawdown data from each monitoring port equipped with a pressure transducer. Data was recorded as early as one minute and as late as 1,600 minutes, depending on the pumping duration and hydraulic responses of each observation ports. An important component of our sampling plan was the concurrent manual collection of hydraulic head data to ensure the correct functioning of pressure transducers. In particular, once the test was started, manual water levels were taken periodically from various ports, primarily targeting ports that had shown large and quick responses. After one and a half hours, manual head measurements were recorded every 30 minutes, with increments increasing to hourly measurements after three hours. Flow rate measurements were recorded every 30 minutes for the first two hours, followed by hourly measurements thereafter. Upon completion of pumping, we monitored the recovery of the hydraulic head levels in each monitoring interval with pressure transducers and water level meters.

We began the pumping tests at PW4-3 PW5-3, PW5-4, and PW5-5 in June of 2013, where we knew from previous investigations (Alexander et al., 2011; Berg and Illman, 2011b, 2013) high pumping rates (>15L/min) can be achieved. We extended the duration of the pumping tests to longer than the previous pumping tests to obtain long-term hydraulic response data.

Subsequently, we performed pumping tests at PW1-3, PW1-4, PW1-5, PW3-2, and PW3-3, where moderate pumping rates (5 - 15L/min) can be achieved (Alexander et al., 2011; Berg and Illman, 2011b, 2013). We also extended the durations of these pumping tests. However, due to significant background water level fluctuations and/or trends, the quality of the data collected during the 2013 pumping tests at these locations was not as good as previous pumping test data.

We then attempted to pump at several locations where low flow rates (2 - 5L/min) were observed previously (i.e., PW2-3, PW3-4, PW3-5), followed by pumping tests at PW1-2, PW1-8, PW2-2, PW3-2, PW4-2, and PW5-2. However, pumping tests at PW1-2, PW1-8, PW2-2, PW3-2, PW3-5, PW4-2, and PW5-2 could not be completed because these locations became dry soon after pumping started, suggesting that the K-values at these locations are low. All extraction pumping tests were completed in September of 2013. Instead of extraction, additional tests at PW1-1, PW3-1, and PW5-1 were conducted in June of 2014 by injecting water at these screen intervals.

A combination of the pumping and injection test data obtained in this study, along with the Berg and Illman (2011b) data, provide a more complete hydraulic response data set for the site. As a result, a total of 15 pumping tests have been conducted at the NCRS; their details are summarized in Table 1. These tests ranged in duration from 4.4 hours to 26.5 hours.

Table 5-1. Summary of Pumping/Injection Tests Performed at NCRS

Well Location	Pumping Rate (L/min)	Duration (hour)	Type
PW1-1	1.89	4.5	Injection
PW1-3	10.50	6	Pumping
PW1-4	6.30	8.5	Pumping
PW1-5	4.40	22.5	Pumping
PW1-6	0.95	6.5	Pumping
PW1-7	1.05	26.5	Pumping
PW2-3	1.91	7	Pumping
PW3-1	0.94	4.4	Injection
PW3–3	2.10	22	Pumping
PW3–4	1.50	22	Pumping
PW4–3	30.20	22.5	Pumping
PW5-1	0.85	4.52	Injection
PW5–3	7.80	22	Pumping
PW5–4	7.80	8.5	Pumping
PW5–5	8.10	22	Pumping

Nearby site activities that may potentially impact the pumping test include the construction of a daycare center approximately 30 meters away from the NCRS well field. The day care center is based on a slab construction and does not have a basement, thus no dewatering activities are taking place. We were monitoring a number of pressure transducers during this construction. Results to date reveal that the impact of the construction on ambient water levels appears to be minimal.

5.4.2 AFP44 Site in Tucson, Arizona

Due to subcontracting issues and project funding distribution delay resulting from the Principal Investigator moving to a different company, the field activities at the AFP44 site were delayed until June of 2014. Using the well inventory provided by the URS Corporation, wells for monitoring were selected based on various planned pumping tests. In June of 2014, in preparation for the first pumping test with the assistance of Ageiss's onsite staff, we were able to locate some of the selected wells and some new wells were found. After we had updated the selection of wells accordingly, we measured the baseline water levels at each selected well and installed the pressure transducers to monitor the background water level changes. From June of 2014 to August of 2015, we installed 44 pressure transducers in an array of selected monitoring wells, including some of the newly discovered wells. The water levels at the wells were continuously monitored and were recorded in two-minute intervals. The recorded groundwater levels were confirmed by independent measurements using a water level sounder. Transducer data were downloaded and the transducers were reprogrammed periodically. Daily average pumping rates at extraction wells and injection wells were reported by URS Corporation (subsequently acquired by AECOM in 2015).

In June of 2014, the pumping system was shut down for around 7 days.

One month later, the pumping rates were reduced by approximately 95 gpm at E-13 and increased by approximately 75 gpm and 25 gpm, at E-8 and E-4, respectively, for several days. In between, the whole system was shut down for one day. The pressure transducers recorded the hydraulic responses and recovery at the selected monitoring wells.

In August of 2014, the pumping at E-6, R-11, R-18, and R-20 were turned off, and the pumping at R-3, R-4, and R-5 were turned on. The pumping rate at E-13 was increased.

In September and October of 2014, drilling and hydraulic fracturing were performed at the AFP44 as a part of the remediation effort. Short-term pumping rate adjustments were made by URS Corporation, and based on operational needs during this time, there were two system shutdowns, one for a few hours and the second one for about a week. However, the transducer data recorded during this period might be affected by the drilling and hydraulic fracturing operations onsite.

In October of 2014, pumping rates at various pumping wells were adjusted by URS Corporation based on operational needs. We were hoping to adjust the pumping rate at E-1. However, due to site operational needs, pumping rate at E-1 could not be changed.

In November of 2014, the system was shut down twice for a few hours and the pumping at E-23 was turned off for several days.

In December of 2014, the complete system was down for approximately one day due to malfunctioning. This event provided an opportunity to obtain the hydraulic response data for the whole system. In addition, pumping at E-24 was shut down for several days. However, hydraulic fracturing at the site was resumed in December, and some of the transducer data might potentially be affected.

From January to March of 2015, pumping rates at various wells could not be adjusted because of the annual sampling event.

In January of 2015, the pumping system was shut down for about three days. One month later, E-23 was turned off twice for a few days.

In March of 2015, pumping rates of a majority of the pumps were reduced or set to zero for short periods.

In April of 2015, the system was turned off for about a week. After the system had resumed, E-24 was shut down again. In the following month, E-6 was turned off and E-24 was turned on instead.

In June of 2015, parts of the system were turned off for a few hours, followed by a system shutdown for two weeks. After the restart in July, E-6 was turned off and the pumping at E-24 resumed.

From the various recorded events, we chose four for the HT interpretation and analysis: Rate change at E-13 in July 2014, the shutdown of E-23 in November 2014, the system shutdown in April 2015 and the rate change at E-24 in May 2015. For validation of the K- and Ss-field generated by HT, the data of the system shutdown in January 2015 were used.

5.5 DATA ANALYSIS AND RESULTS

5.5.1 NCRS at UW, Canada

The NCRS data selected for this study includes potentiometric head responses recorded as early as one minute and as late as 1,600 minutes after commencing, depending on the pumping duration and hydraulic responses of each observation ports. One to three points were selected to define each curve. During some pumping tests, there were some ports showing negligible responses. These data were also included in the inversion as they provide information regarding lack of hydraulic connectivity between the pumped and observation ports. Seven pumping tests (PW1-1, PW1-4, PW1-6, PW2-3, PW3-3, PW4-3, and PW5-3) are used for calibration, while the other seven pumping tests (PW1-3, PW1-5, PW1-7, PW3-1, PW3-4, and PW5-5) are selected for model validation purposes. Pumping tests used for calibration encompass the top and bottom pumping ports, as well as four corner wells in the central 15 m x 15 m pumping and observation area, to provide more spatially different flow information for HT analysis. In total, 195 pressure head data values were selected for model calibration, and 176 head data values were used for model validation.

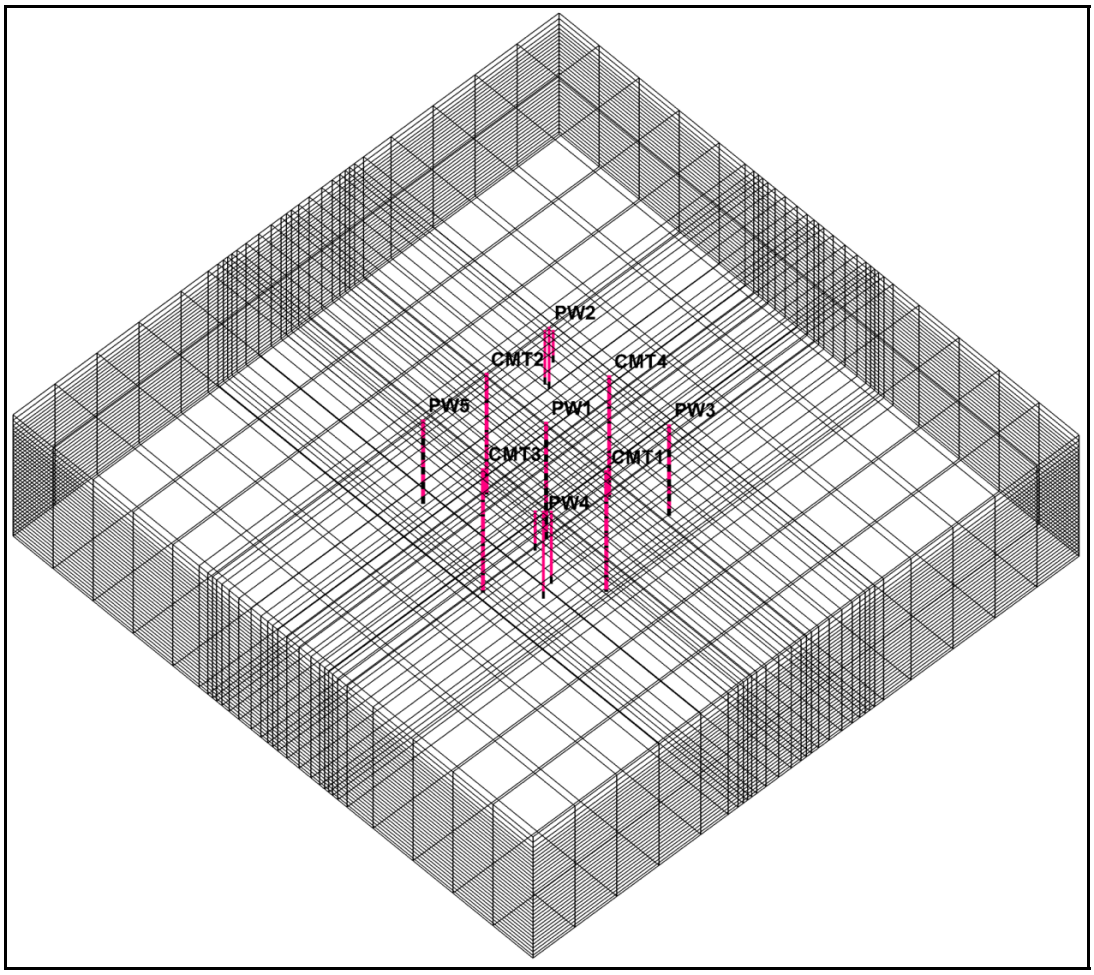


Figure 5-6. Computational Grid for NCRS

In order to simulate groundwater flow for both forward and inverse modeling, a three-dimensional domain of 70 m × 70 m × 17 m was discretized into 31,713 computational elements with 34,816 nodes (Figure 5-6). This grid is similar to the one used previously by Berg and Illman (2011b, 2013, 2015) in terms of the general layout, but has slightly larger domain sizes to include more wells with known borehole data. The elements are gradually refined from the boundary areas to the vicinity of central well locations, decreasing from grid block sizes of 5 m × 5 m × 0.5 m to 0.5 m × 0.5 m × 0.5 m. All groundwater flow simulations are conducted using the finite element code MMOC3 (Yeh et al. 1993). For boundary conditions, the top and bottom faces are defined as no-flow boundaries, while the other four faces are kept as constant head boundaries, as in the previous studies by Berg and Illman (2011b, 2013).

Three different parameterization cases (referred to as Case 1 through Case 3) were considered for inverting the HT data in this study: (1) an effective value approach by treating the model as homogeneous, (2) a zonation approach based on geological stratigraphy, and (3) a highly parameterized geostatistical approach. HT inversion was performed using the Simultaneous Successive Linear Estimator (SimSLE) code, which can invert all the data sets simultaneously, thus providing more constraints to the inverse problem (Xiang et al., 2009) compared to when the data are sequentially included in the inverse code (Yeh and Liu, 2000; Berg and Illman, 2013).

In SimSLE, natural log values of a hydraulic conductivity (i.e., ln K) in the heterogeneous field are treated as a stochastic process, and the stochastic conditional means of these parameters are used for groundwater flow modeling in the aquifer. Conventional model calibration was performed using the commonly used parameter estimation software PEST. To quantitatively assess the performance of model calibration and validation results of all inversion models, the mean absolute error (L_1) and mean square error (L_2), were calculated as:

$$L_1 = \frac{1}{n}\sum_{i=1}^{n}|h_i^* - h_i|$$
$$L_2 = \frac{1}{n}\sum_{i=1}^{n}(h_i^* - h_i)^2 \quad (5\text{-}1)$$

where n is the total number of pressure heads used for calibration, h_i is the i^{th} observation head data, and h_i^* is the corresponding simulated head.

Case 1: Effective Parameter Approach

We considered two cases (Case 1a and Case 1b) in the effective parameter approach. Case 1a treats the aquifer to be isotropic, where we estimate only K_{eff}, and Case 1b treats the entire simulation domain to be anisotropic, for which we estimate the effective K_x, K_y and K_z. An initial value of 8.0×10^{-6} m/s with a minimum bound of 1.0×10^{-9} m/s and a maximum bound of 1.0×10^{0} m/s were used for PEST calibration. The initial value that we chose is the geometric mean of individual K estimates obtained by matching the transient drawdown curve at each observation port during pumping at PW1-3 well by Berg and Illman (2011b).

The simultaneous calibration of the effective parameter model with data from seven pumping tests for the isotropic Case 1a yielded an estimated K_{eff} of 8.4×10^{-6} m/s and a corresponding uncertainty indicated by the 95% confidence interval, which has an upper limit of 9.8×10^{-6} m/s and lower limit of 7.2×10^{-6} m/s. For the anisotropic Case 1b, K_x was estimated as 1.04×10^{-5} m/s with an upper limit of 1.54×10^{-5} m/s and a lower limit of 7.02×10^{-6} m/s, and K_y was estimated as 1.19×10^{-5} m/s with an upper limit of 1.68×10^{-5} m/s and a lower limit of 8.36×10^{-6} m/s. The effective K values in the horizontal directions x and y are similar. The value of K_z was lower than K_x and K_y, estimated as 6.37×10^{-7} m/s with an upper limit of 1.08×10^{-6} m/s and a lower limit of 3.75×10^{-7} m/s.

When treating the heterogeneous aquifer to be uniform, the estimated parameters are found to be dependent on the observations, as well as pumping locations (e.g., Wu et al. 2005; Straface et al. 2007; Wen et al. 2010, Huang et al. 2011; Sun et al. 2013; Berg and Illman 2013; 2015). The previous transient HT study by Berg and Illman (2015) found that the effective parameters varied depending on the location of pumping tests when estimating these values for each pumping test at NCRS. Therefore, the estimated effective K values from Case 1a and 1b should be more representative of the test area in an average sense, since the effective K values are estimated by simultaneously considering data from all seven pumping tests.

Case 2: Geological Zonation Approach

Two geological models were constructed and calibrated: the 5-layer (Case 2a) and the 19-layer (Case 2b) models. The latter is shown in Figure 4-4. Both geological models were discretized using the same grid as in the other cases. K values of the elements located in the same layer were treated to be uniform and isotropic. The initial K value for calibration of the 5-layer geological model was also set as 8.0×10^{-6} m/s with a minimum bound of 1.0×10^{-9} m/s and a maximum bound of 1.0×10^{0} m/s. For the 19-layer model, however, the estimated K values of Case 2a were used as initial values for PEST calibration, due to the difficulty in obtaining enough observable pumping test data from low permeability Clay/Site and Clay layers located at bottom layers while pumping from wells located in the top layers (PW1-1, PW1-3, PW3-1, etc.) of the aquifer-aquitard system. Previously, through Bayesian model analysis in a sandbox aquifer, Schöniger et al. (2015) suggested that a lower-complexity geological model will be more likely to be justified than the higher-complexity models with a given amount of observed pressure data. Zhao et al. (2016) also proved that a simplified geological model with fewer layers leads to consistent estimates of K values when observation density is lowered. Thus, it is reasonable and convenient for calibrating the 19-layer geological model while the results of a 5-layer geological model are used as initial values for current HT study.

For this case, the simultaneous calibration of the 5-layer geological model (Case 2a) was completed after 172 model calls. The estimated K values and the corresponding 95% confidence intervals and the layer definitions are listed in Table 5-2, while the estimated K distribution is presented in Figure 5-9a. Generally, the calibration of the 5-layer geological model yielded the highest K value for the sand and gravel layer (layer 15) and the lowest K value for the bottom merged layer 16^{*}, consisting of silt and clay layers (layer 16 through 19). K estimates for merged layer 1^{*} and 12^{*} are close to the initial value of 8.0×10^{-6} m/s, which may be the result of using a single layer for multiple soil types. In addition, the upper sand layer (layer 11; Figure 4-4) known to have a high K value, was assigned a value of 7.74×10^{-8} m/s, suggesting that the layer is a low K zone, which is inconsistent with known geological information.

The estimated K values for layers 3 and 5 have significantly large 95% confidence intervals comparing to those of the other layers. One reason is that layers 3 and 5 only exist in a narrow portion of the geology model and also are far from the pumping and observation wells, as shown in Figure 4-4, thus very few, or no observation data are available in these layers to provide the pressure head information for model calibration. Similar results are found in Zhao et al. (2016) through laboratory sandbox study where the geological zonation information is perfectly known.

Comparing the results Table 5-3 to Figure 5-9a, K estimates for the main sand layers of the 19-layer model show some similarities to the 5-layer geological model, by estimating a relatively high K value for Layer 15, while estimating a low K value for Layer 11. Differences between the two geological models are obvious from Figure 5-9a and Figure 5-9b. Firstly, more details about the interlayering of high and low permeability zones are revealed in Case 2b for the upper part of the domain than in Case 2a. Secondly, variations of K estimates are introduced for low permeable layers (layer 16 to layer 19) at the bottom of the study area.

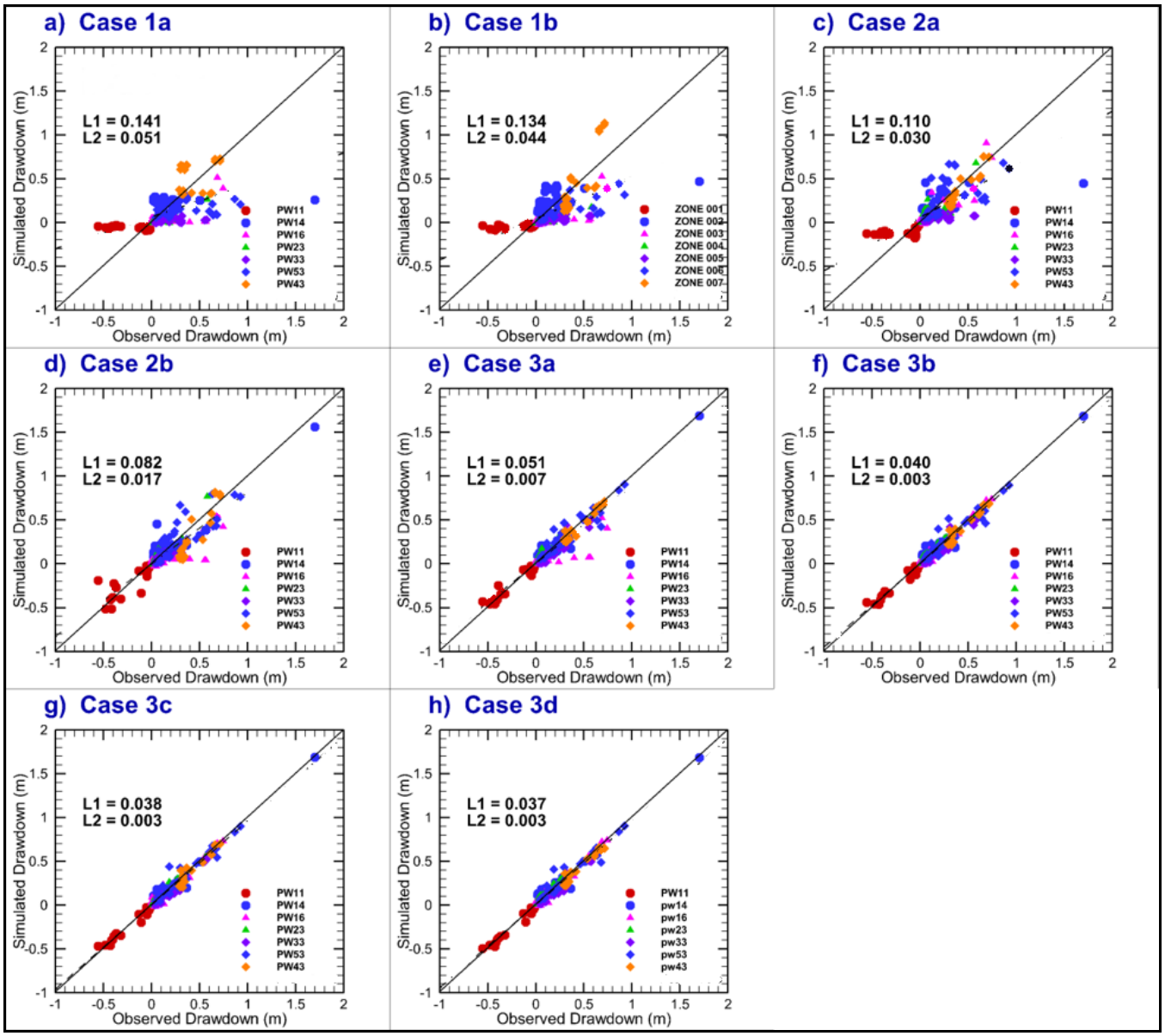


Figure 5-7. Scatterplots of Observed Versus Simulated Drawdowns for Model Calibrations Using Seven Pumping Tests ¶for the: (a) Isotropic Effective Parameter Model; (b) Anisotropic Effective Parameter Model; (c) Geological Model with Five Layers; (d) Geological Model with 19 Layers; (e) SimSLE starting with K = 8.0 × 10-6 m/s as prior mean; (f) SimSLE using the Calibrated Five-layer Geological Model as Prior Distribution; (g) SimSLE Using the Calibrated 19-layer Geological Model as Prior Distribution; and (h) SimSLE Using the Uncalibrated 19-layer Geological Model Assigned with Permeameter K Values as Prior Distribution.

The solid line is a 1:1 line indicating a perfect match.

Case 3: Geostatistical Inversion Approach

The inversion process starts with cokriging using available measurements of hydraulic property and pressure heads to produce the conditional property field, with the assumptions that the unconditional means, spatial covariance functions and structure parameters (correlation scales λ_x, λ_y, λ_z and the variance, $\sigma_{\ln K}^2$) of hydraulic parameters are known. In this study, the exponential covariance model is adopted for the estimated parameter field. The initial guesses for correlation scales of the K-field are set as $\lambda_x = \lambda_y = 4$m, and $\lambda_z = 0.5$m, and a variance is set to be $\sigma_{\ln K}^2 = 5$, which are kept at the same values used in Berg and Illman (2011b). The cokriged parameter field is then iteratively updated by SimSLE to minimize the differences between observed and simulated heads.

Four scenarios (Case 3a, 3b, 3c and 3d) are considered using different prior distributions. In Case 3a, the inversion starts with a uniform mean field of $K = 8.0 \times 10^{-6}$ m/s, which is the same as the initial K value used in the effective parameter and geological zonation approaches. On the other hand, for the other three cases (Cases 3b – 3d), geologic information is used as prior knowledge for the inversion. Specifically, Case 3b used the estimated K values from Case 2a as the prior mean distribution; Case 3c used the K estimates from Case 2b as the prior mean distribution; Case 3d used the 19-layer geological model (Case 2b) populated with permeameter tested K values as the prior mean distribution. In Case 3d, the corresponding K values for each layer were calculated as the geometric mean of soil sample measurements located in the same layer, and these values are listed in Table 5-3. In Table 5-3, permeameter test K values of layer 5 and layer 19 are estimated to be the same as the values of layers 6 and 17, respectively, due to the fact that no core samples are available for layers 5 and 19, but having similar soil material with layer 6 and 17.

Through Case 3b, 3c and 3d, we test the impact of using both calibrated geological models and permeameter test K values as prior mean K distributions for the geostatistical inversion approach. Thus the findings of Case 3b and 3c would be more useful for practitioners than Case 3d, since using detailed permeameter test results will need additional efforts.

The L_2 norm changes during the calibration process for all four scenarios are plotted in Figure 5-8. We selected inversion results from the iteration step at which the L_2 norm has stabilized, indicating the convergence of the inversion process as suggested by Xiang et al. (2009). The result from the 82nd iteration was selected for Case 3a, while results from the 62nd iteration were selected for Cases 3b, 3c and 3d (Figure 5-8). It is interesting to note that Case 3a with a uniform K value takes more iterations to converge when compared to Cases 3b – 3d, in which various geological models are used as prior distributions.

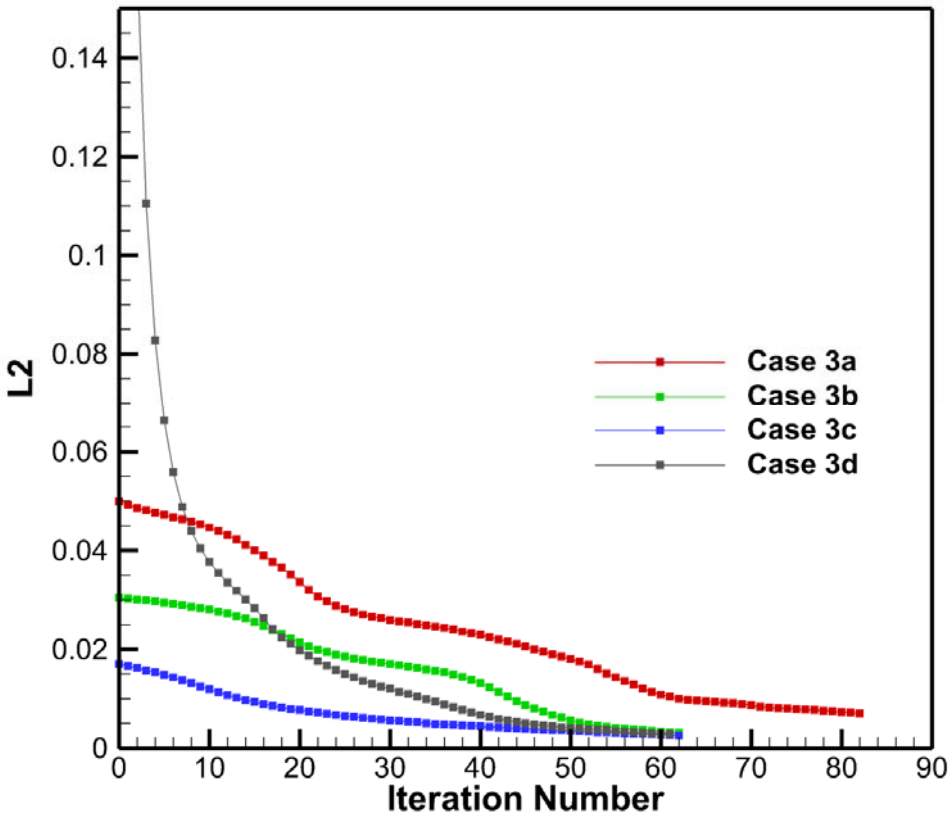


Figure 5-8. Mean Square Error (L_2 norm) as a Function of Iteration Number

For Case 3a, a uniform mean K-field was used as the prior distribution for the geostatistical inversion before SimSLE started to condition the parameter field with pressure head measurements iteratively. The drawdown scatterplot of Case 3a (Figure 5-7e) showed significant improvement over the effective parameter (Case 1a and 1b) and the geological zonation (Cases 2a and 2b) approaches in terms of the L_1 and L_2 norms. However, we see that there is an obvious drift in the data from PW1-6 from the 45-degree line compared to those from the other pumping tests.

Figure 5-9c provides the estimated K tomogram for Case 3a, while its corresponding residual variance of ln K in Figure 5-10a. Examination of Figure 5-9c reveals that, in general, the interlayer patterns of the high and low permeable zones are captured in the central part of modeling domain, where lnK residual variances (Figure 5-10a) are lower. In addition, a higher K zone is visible in the bottom left portion of the domain.

In the previous work of Berg and Illman (2013), who utilized four pumping tests (PW4-3, PW1-3, PW5-3 and PW3-3) to conduct SSHT analysis of the same area, the entire bottom area of the central model domain was estimated to have high K values despite the fact that core samples indicated the presence of low K silt and clay layers. In addition, the lowermost ports situated in the low K zones did not yield measurable drawdown responses during those tests.

In contrast, for this study, we obtained measurable drawdown responses from the bottom observation ports by pumping from port PW1-6 (Figure 5-3) located in the lower part of the domain. By including these additional drawdown data from the low permeable zone, the inversion of all tests yielded slightly improved K estimates without showing the entire bottom area as a high K zone (Berg and Illman, 2011b, 2013). However, our results are still inconsistent with the known geology consisting of silt and clay layers. For example, in the bottom left portion of the domain beyond the central 15m x 15m well cluster area, the K estimates (Figure 5-9c) and the residual variances of lnK (Figure 5-10a) are generally high, due to the fact that no wells and pressure head data are available in that region for model calibration.

In Case 3b, we extended Case 3a by using the K tomogram obtained from Case 2a, which is the calibrated 5-layer geological model as the prior distribution for the geostatistical inverse model. The drawdown scatterplot of Case 3b (Figure 5-7f) reveals an obvious improvement compared to Case 3a (Figure 5-7f e), in which a uniform mean K-field is used as the prior distribution. On the other hand, obvious differences can be seen in the estimated K tomogram from Case 3b (Figure 5-9d) when compared to Case 3a (Figure 5-9c). Specifically, K estimates from Case 3b (Figure 5-9d) reveal a pattern that preserves the geological features of the K distribution of the calibrated 5-layer model, as well as the heterogeneity features in the upper part of the K tomogram for Case 3a (Figure 5-9).

For the bottom part of the simulation domain, the estimated K values in Case 3b (Figure 5-9d) are significantly lower than Case 3a (Figure 5-9c). In addition, the low K zone at the bottom of the simulation domain extends across the site. Both of these features in Case 3d are more consistent with our knowledge of site geology.

The residual variance of ln K for Case 3b (Figure 5-10b) reveals that the variances are relatively low within and in the vicinity of the well field. However, as in Case 3a, the variances are higher away from the well field.

The geostatistical inverse modeling of Case 3c was performed by using the K estimates of the calibrated 19-layer geological model of Case 2b as the prior distribution. From Figure 5-7 we observe that the calibration scatterplot between simulated and observed drawdowns for Case 3c (Figure 5-7 g) shows slight improvements compared to Cases 3a (Figure 5-7e) and 3b (Figure 5-7f). The estimated K tomogram and the corresponding ln K residual variance are provided in Figure 5-9e and Figure 5-10c, respectively. Generally, the main layering pattern shown in Figure 5-9e follows the pattern of the calibrated 19-layer geological model (Figure 5-9b). In addition, we can clearly see more details to the geological features throughout the site, because a 19-layer model is used as the prior distribution.

Case 3d uses the 19-layer geological model populated with permeameter test K values as a prior distribution for geostatistical inverse modeling. This case could be viewed as the scenario with most data included into the inverse model among all four geostatistical inversion cases that include pressure heads, geological data, and local K data from permeameter tests.

Figure 5-7h provides the drawdown scatterplot for Case 3d which shows that the L_1 and L_2 norms have improved over Case 3a. However, the results are comparable to Cases 3b (Figure 5-7f) and 3c (Figure 5-7e) suggesting that including permeameter K data as prior information has not significantly improved the calibration results.

Figure 5-9f provides the estimated K tomogram from Case 3d and the corresponding lnK variance in Figure 5-10d. Compared to the K tomogram for Case 3a (Figure 5-9c), the structural features shown in the geological model (Figure 4-4) are better preserved in the recovered K tomogram (Figure 5-9f). Similar to Cases 3b (Figure 5-9d) and 3c (Figure 5-9e), the heterogeneous K distributions and in particular, the layering is similar to prior values. However, due to the inclusion of permeameter K data into the Case 3d model, the K values for the lower most layer consisting of silt and clay are more representative of site geology in comparison to Cases 3a (Figure 5-9c), 3b (Figure 5-9d), and 3c (Figure 5-9e). The residual variance map of lnK (Figure 5-10d) is similar to the other cases (Figure 5-10a–c).

Results obtained from calibrating Cases 3b, 3c and 3d suggested that when geologically distributed K-fields are used as prior distributions, HT analysis using the geostatistical inversion approach could yield K tomograms with geological features. This would be helpful for HT to correctly capture the stratigraphic features for areas where only limited pressure head data are available.

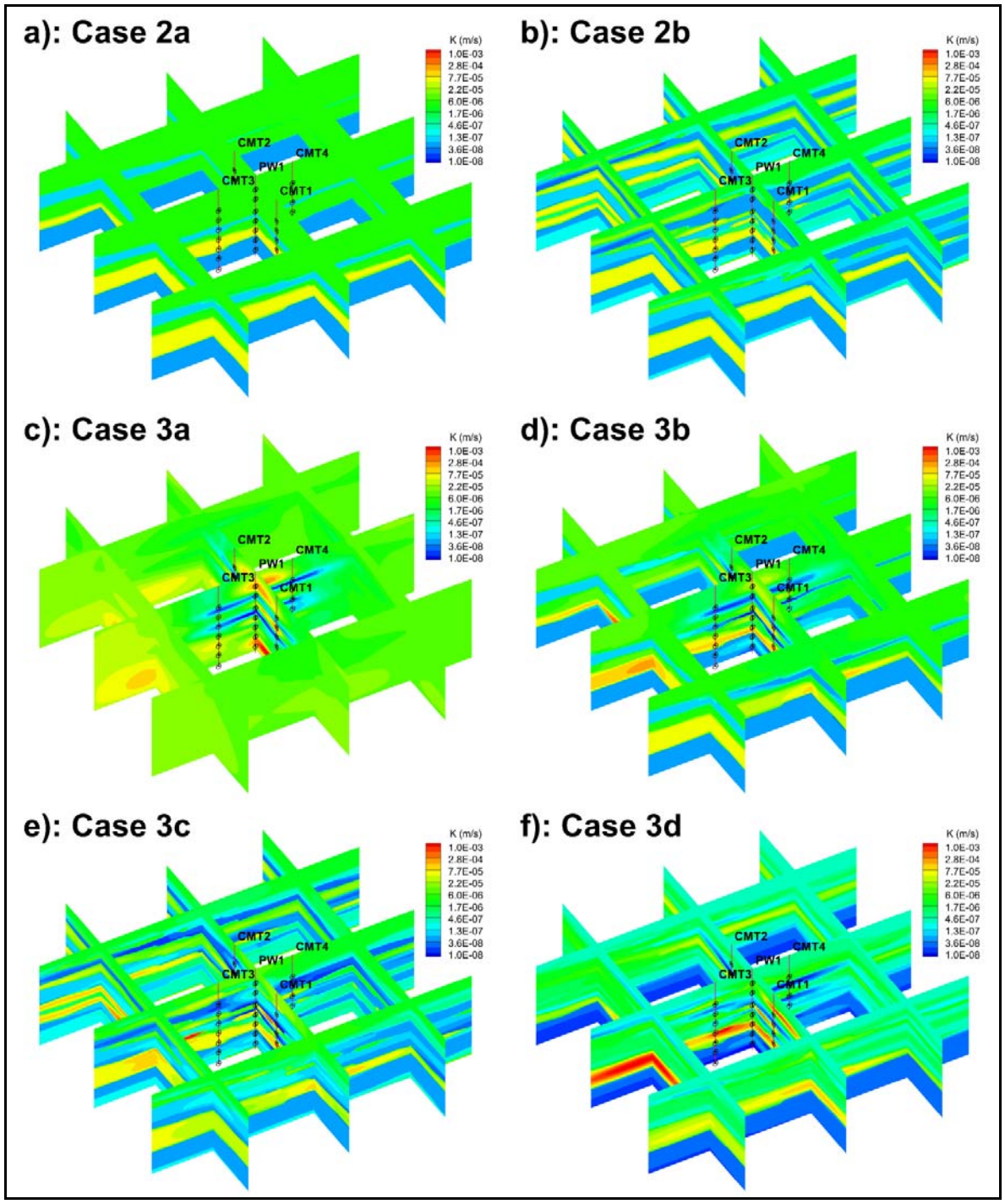


Figure 5-9. Estimated K-fields from the Inversion of Seven Pumping Tests ¶for: (a) the Geological Model with Five Layers; (b) the Geological Model with 19 Layers; (c) SimSLE Starting with a Uniform K = 8.0 × 10-6 m/s; (d) SimSLE Using the Calibrated Five-layer Geological Model as Prior Distribution; e) SimSLE Using the Calibrated 19-layer Geological Model as Prior Distribution; (f) SimSLE Using the Uncalibrated 19-layer Geological Model Assigned with Permeameter Test K Values for Each Layer as Prior Distribution.

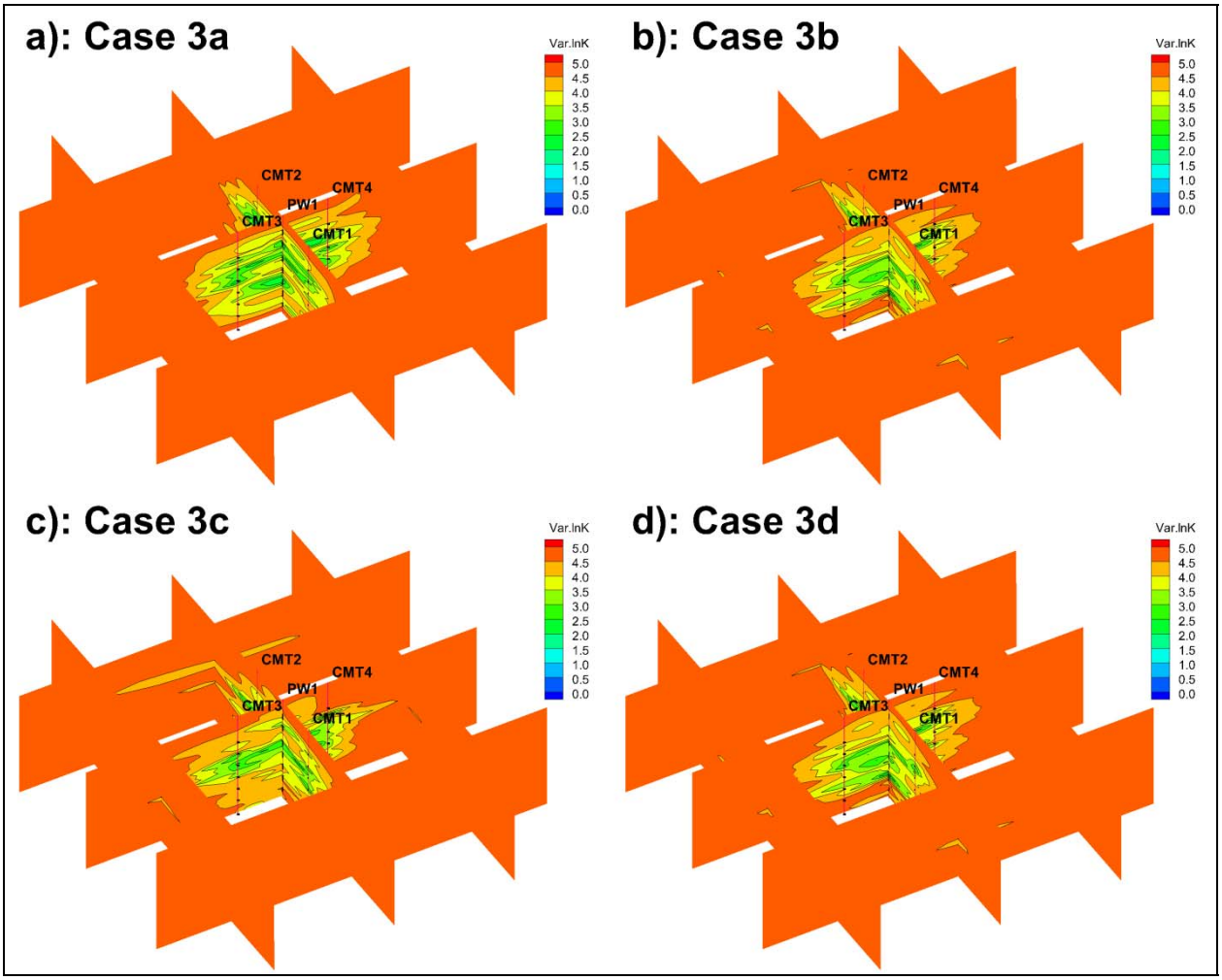


Figure 5-10. Corresponding Residual Variances of Estimated lnK-fields from the Inversion of Seven Pumping Tests ¶or (a) Case 3a: SimSLE Starting with a Uniform K = 8.0×10^{-6} m/s; (b) Case 3b: SimSLE Using the Calibrated 5-layer Geological Model as Prior Distribution; (c) Case 3c: SimSLE Using the Calibrated 19-layer Geological Model as Prior Distribution; and (d) Case 3d: SimSLE Using the Uncalibrated 19-layer Geological Model Assigned with Permeameter Test K Values for Each Layer as Prior Distribution.

Table 5-2. Estimated K values and Corresponding Posterior 95% Confidence Intervals for the 5-layer Geological Model

Layer	Estimated K (m/s)	95% Confidence Intervals	
		Lower limit	Upper limit
1*[a]	5.33×10^{-6}	4.22×10^{-6}	6.74×10^{-6}
11	7.74×10^{-8}	3.94×10^{-8}	1.52×10^{-7}
12*[b]	5.12×10^{-6}	4.47×10^{-7}	5.87×10^{-5}
15	6.38×10^{-5}	4.63×10^{-5}	8.78×10^{-5}
16*[c]	4.84×10^{-8}	3.01×10^{-8}	7.80×10^{-8}

[a] Layer 1* is a merged layer of the original Layers 1 through 10;

[b] Layer 12* is a merged layer of the original Layers 12 through 14;

[c] Layer 16* is a merged layer of the original Layers 16 through 19.

Table 5-3. Soil Type, Permeameter Test K, Estimated K Values and Corresponding Posterior 95% Confidence Interval Limits for the 19-layer Geological Model

Layer	Soil Type	Permeameter Test K (m/s)	Estimated *K* (m/s)	95% Confidence Intervals	
				Lower limit	Upper limit
1	Clay	2.65×10^{-7}	3.15×10^{-6}	2.88×10^{-7}	3.44×10^{-5}
2	Silt	7.76×10^{-7}	4.26×10^{-7}	5.64×10^{-15}	$3.21\times10^{+1}$
3	Sand	1.45×10^{-7}	5.83×10^{-7}	1.11×10^{-19}	$3.07\times10^{+6}$
4	Clay	6.09×10^{-8}	1.27×10^{-5}	1.32×10^{-7}	1.23×10^{-3}
5	Sand and Silt	2.38×10^{-6} [a]	8.51×10^{-8}	8.41×10^{-31}	$8.60\times10^{+15}$
6	Sandy Silt	2.38×10^{-6}	3.53×10^{-8}	2.50×10^{-9}	4.97×10^{-7}
7	Silt	3.13×10^{-7}	4.01×10^{-8}	1.34×10^{-8}	1.20×10^{-7}
8	Clay	1.82×10^{-6}	1.35×10^{-5}	2.86×10^{-6}	6.36×10^{-5}
9	Sandy Silt	5.04×10^{-6}	5.40×10^{-5}	2.14×10^{-5}	1.36×10^{-4}
10	Silt	7.47×10^{-6}	2.34×10^{-5}	4.50×10^{-6}	1.22×10^{-4}
11	Sand	1.32×10^{-6}	1.05×10^{-7}	4.90×10^{-8}	2.23×10^{-7}
12	Clay	3.74×10^{-7}	3.66×10^{-8}	5.54×10^{-9}	2.42×10^{-7}
13	Sandy Silt	1.17×10^{-6}	6.29×10^{-5}	3.15×10^{-5}	1.25×10^{-4}
14	Silt	1.13×10^{-7}	6.27×10^{-7}	2.82×10^{-7}	1.39×10^{-6}
15	Sand and Gravel	1.22×10^{-5}	6.66×10^{-5}	5.31×10^{-5}	8.35×10^{-5}
16	Clay	2.01×10^{-8}	5.84×10^{-8}	2.32×10^{-9}	1.47×10^{-6}
17	Silt and Clay	2.44×10^{-8}	4.18×10^{-7}	2.95×10^{-11}	5.94×10^{-3}
18	Clay	4.72×10^{-9}	2.37×10^{-7}	6.16×10^{-13}	9.15×10^{-2}
19	Silt and Clay	2.44×10^{-8} [b]	1.70×10^{-5}	8.49×10^{-8}	3.41×10^{-3}

[a] K value for layer 5 is estimated as the value of layer 6;

[b] K value for layer 19 is estimated as the value of layer 17.

5.5.2 AFP44 Site in Tucson, Arizona

Identification of pumping/injection perturbation events from pumping data

Injection and pumping rates at the AFP-44 pump-and-treat are obtained from the site contractors and include daily injection or pumping at each of the wells. An overview of the injection and extraction rates of each well and the percentage of the total system injection and extraction is given in Table 5-4. Examination of pumping and injection records indicate that there are three main types of events contributing to head changes: (i.) system shutdown and resume, (ii.) changes in pumping and injection strategy, and (iii.) fluctuation of flow rates (see Figure 5-11). The injection records appear to be quite smooth; the site does not have any storage facility for treated groundwater, so the total injection rate of the system must equal total pumping rate (plus loss from leaks in the pipeline).

We have observed and recorded 10 system shutdown events, the longest of which occurred on June 12-28, 2015. There are also several notable shutdown and recovery events of individual wells. For example, E-13 reduced its pumping from ~300 GPM to ~100GPM in mid-July 2014. Also, the site management adopted a new injection strategy in mid-August 2014 such that R11, R18, and R20 are turned off while R3, R4, and R5 are turned on. The pumping rates among E4, E5, and E6, as well as between E23 and E24 seem to be interdependent, and each of the groups is a cluster of wells next to each other. E2 was turned on in late August 2015, towards the end of our record (not shown in Figure 5-11). Finally, E1, the pumping well located in the center of the site, was turned off until July 7, 2015 when system resumed from shutdown on June 30, 2015.

Table 5-4. The Pumping and Injection Rates for the Extraction and Injection Wells (based on the record of June 15, 2014)

The relative percentages that exceed 10% are highlighted.

Recharge Well	**R2**	**R3**	**R4**	**R5**	**R8**	**R9**	**R10**	**R11**	**R14**	**R18**	**R20**
Rate (m^3/d)	2180	55	55	55	1581	1744	491	1962	55	436	545
Percentage (%)	***23.8***	0.6	0.6	0.6	***17.3***	***19.0***	5.4	***21.4***	0.6	4.8	6.0
Extraction Well	**E1**	**E2**	**E4**	**E5**	**E6**	**E7**	**E8**	**E9**	**E13**	**E23**	**E24**
Rate (m^3/d)	-1145	-55	-654	-55	-600	-1139	-709	-600	-1690	-1090	-1145
Percentage (%)	***12.9***	0.6	7.4	0.6	6.8	***12.8***	8.0	6.8	***19.0***	***12.3***	***12.9***

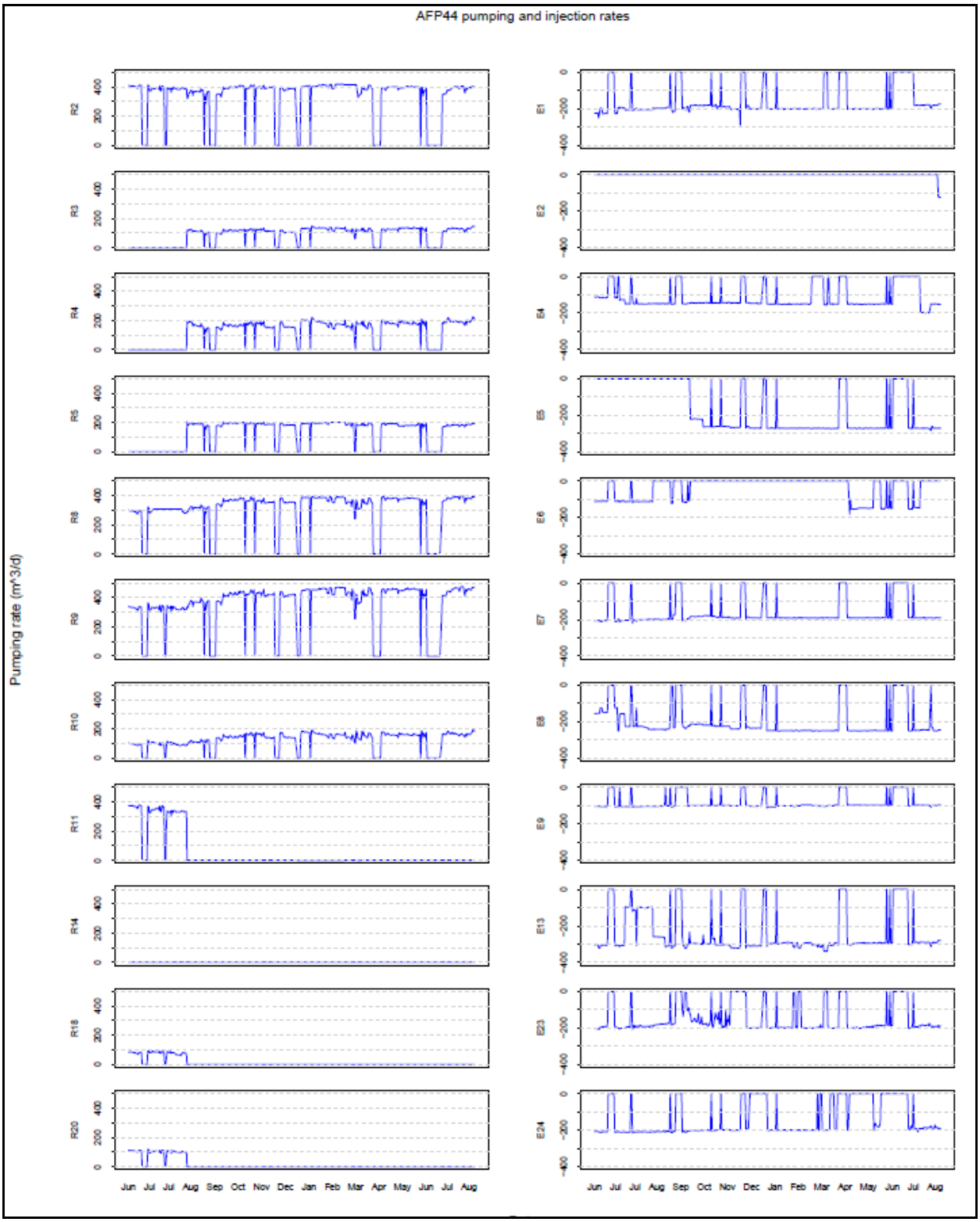


Figure 5-11. Time Series of Pumping Rates

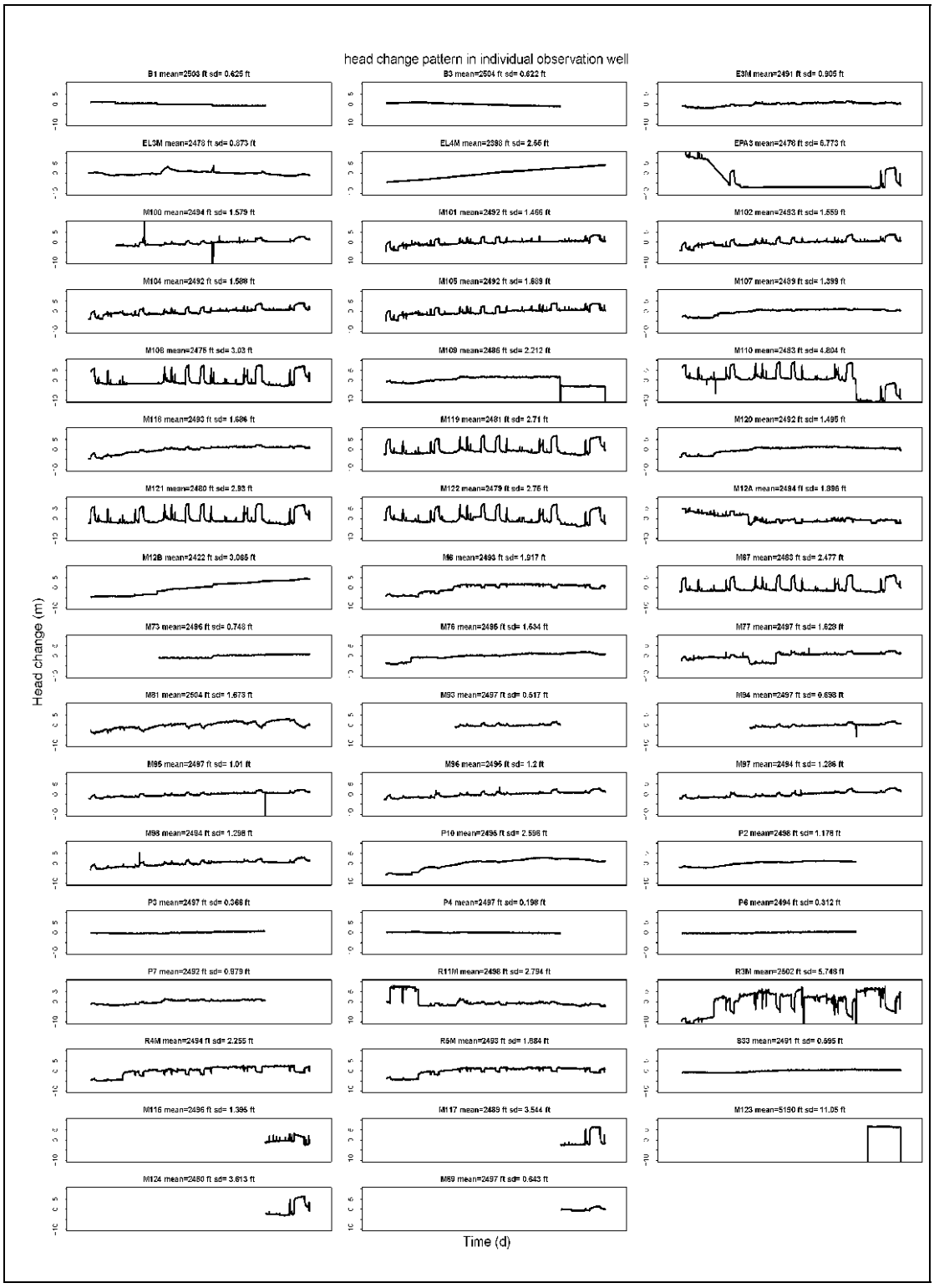


Figure 5-12. Individual Variations of Head During the Monitoring Period (June 2014 to July 2015)

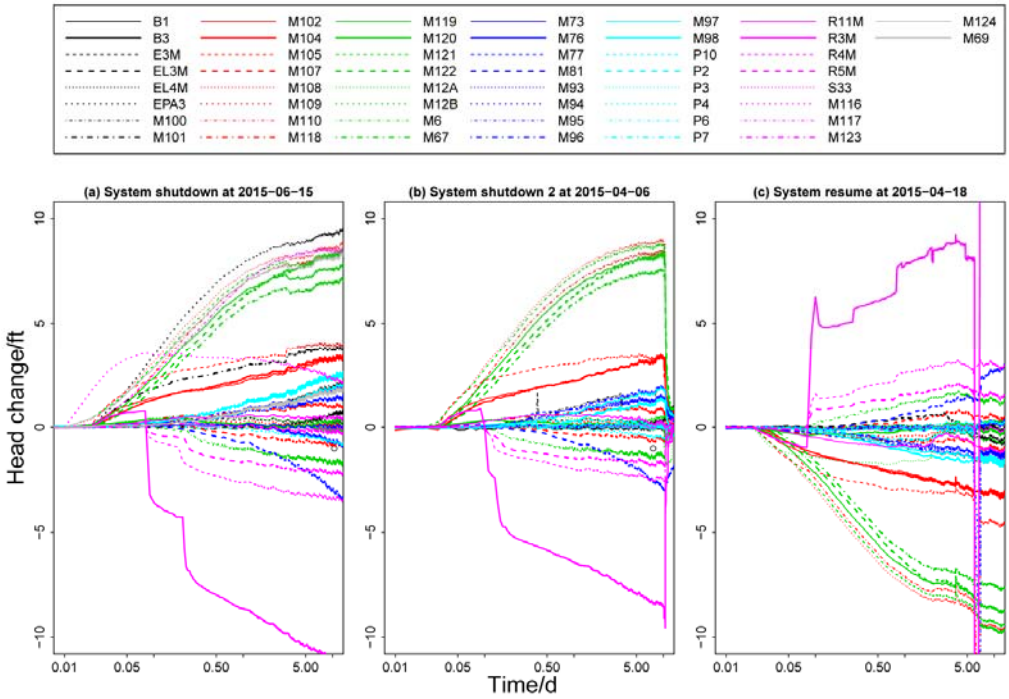
Mean is the mean head elevation in this well; sd is the standard derivation of the head. Generally larger sd value means that this well is more active to the change of stress change (injection or pumping).

We have little knowledge about flow rate changes within the day, and we have to assume they have minimal impact on parameter estimation. For turning wells on and off, which is the most typical type of flow rate changes at pump-and-treat sites, the above has a minor impact because if flow rates before and after the change are known, it is straightforward to determine the time of the change from hydrograph responses.

Another source of uncertainty in flow rates stems from the integrity of wells. The two most powerful wells in the pump-and-treat systems, E-13 (formerly HAC-1) and R-2 (formerly HAC-3), were built in 1952 and 1954, respectively, as water supply wells. They were originally drilled to 600 ft. and 400 ft. below land surface, respectively. Although their screen intervals below 250 ft. were abandoned after TCE was discovered at the site as a precaution to prevent TCE from migrating to the regional aquifer, it remains plausible that water from the deeper, abandoned intervals contributed to the high productivity of these two wells.

Evaluation of hydraulic responses to pumping/injection perturbation events

Head records from observation wells reveal significant layering at the site (Figure 5-12). Heads in wells screened in the Shallow Groundwater Zone (SGZ, i.e. above UZUU) and the Lower Zone do not change with time, indicating they do not respond to changes in pumping or injection in UZ. In the well network, some wells form pairs that screen at UZUU and UZLU, respectively at the same horizontal location. The response within each of the pairs varies greatly, indicating the aquitard that subdivides UZ is an effective flow barrier. It seems that observation wells screened across UZUU are more responsive to injection changes, while those screened across UZLU are more responsive to pumping changes, as most injection wells are screened across UZUU and most pumping wells are screened across UZLU. The extent of the UZ pinchout to the west of the site, however, is not clearly revealed from visual examination of data.



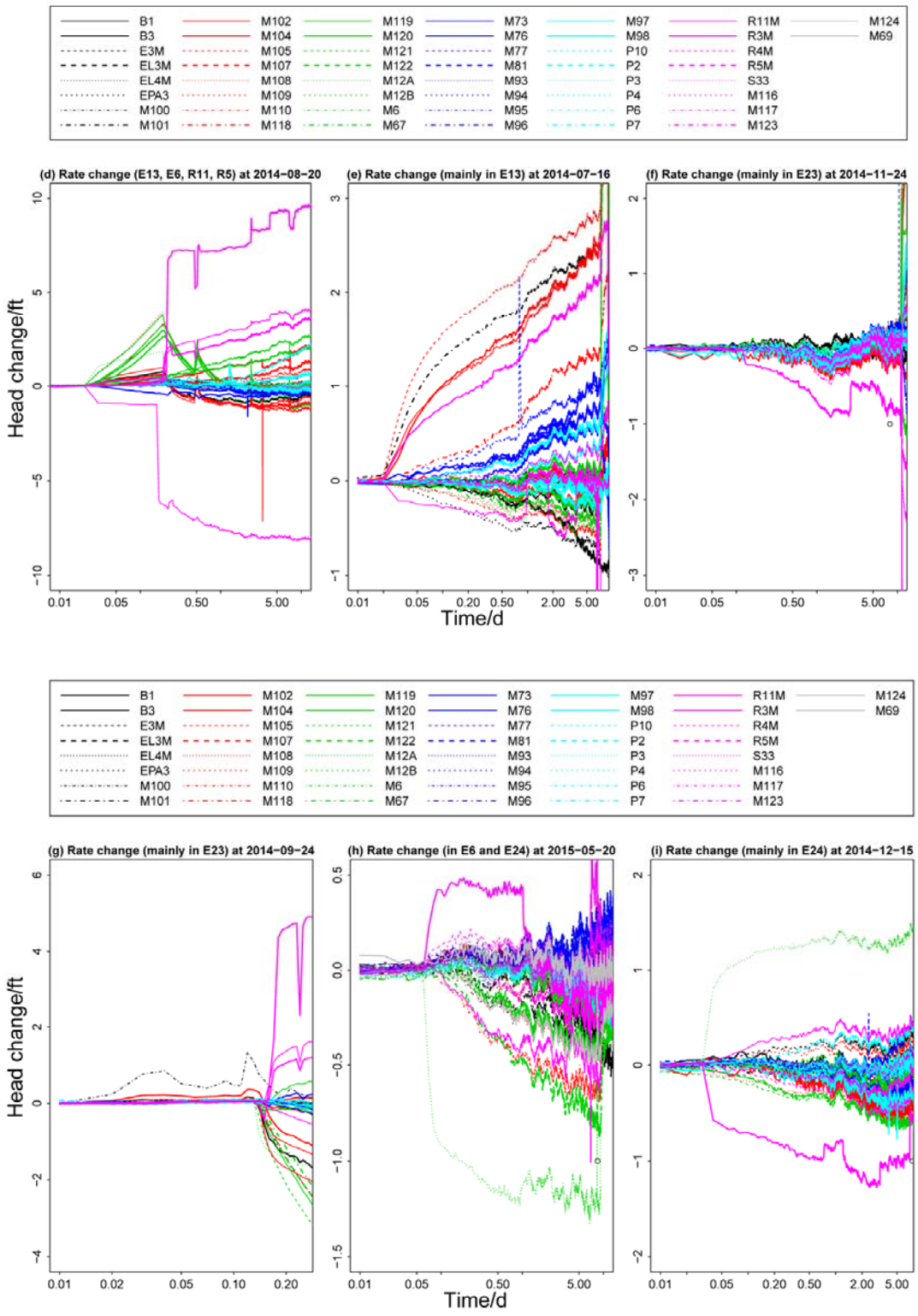


Figure 5-13. Head Changes in Observation Wells during Several Events

b, e, f, and h are selected in the inversion based on manual check. The other events are not used due to either redundancy (many system shutdown events) or unclear/ unreasonable responses.

Figure 5-13 lists the detailed head changes in observation wells during some events. The head changes are differences between heads before and after events. In most situations, the heads are constants between two events; meaning quasi-steady state flows are fully established. Since saturated groundwater flow is a linear problem, the principle of superposition can be used and each event can be treated as an independent pumping test. Figure 5-13(a) and Figure 5-13(b) represent selected system shutdown events. As shown, during these shutdown events, the head changes show very similar patterns. That means the ten shutdown events are consistent and they can be integrated to reduce the effect of observation noise. By examining individual drawdown/buildup curves, we find that the heads at recharge wells (R3M, R4M, R5M, etc.) show rapid decreases in pressure due to the shutdown of these recharge wells. M81, M6, and M107 show mild decreases, meaning they are more influenced by recharge wells than by pumping wells. This is reasonable since they are close to the recharge wells. On the other hand, M108, M121, M110, M119, M122, and M67 have a significant increase of head during the shutdown, which means these wells are well-connected to the pumping wells. For instance, the above wells are all close to E-2, E-1, and E-13 (excluding M-67, which has a long screen interval). M105, M104, M102, M100, M96, M97, M98, M93, M94, and M95 are observed to have a moderate increase of heads. Some of the observation wells are not sensitive to the shutdown (and other events), like B1, B3, R11M, M69, S33, M73, M76, M120, M12B, E-3M, EL4M, M107, M109, P2-P7, and P10. The system resumption event in Figure 5-12(c) also confirms that those shutdown-recovery events are consistent; the head changes in resumption events are identical in shutdown events, but inversely so. During these shutdown-resume events, only one or two shutdown events are necessary for the hydraulic tomography inversion.

In middle August, 2014, R11, R20, and R18 were shut down, and R3, R4, and R5 were turned on. Also, E6 was shut down, and E13 increased the pumping rates. Due to these changes, the head changes experienced are shown in Figure 5-13(d). We find that R3M and R11M had significant head increase and decrease. R4M and R5M also had head increase, which can be attributed to the resumption of R4 and R5. It is also shown that the response dates are not exactly the same as the record of pumping/injection rates (Figure 5-11, Figure 5-12). The head response data were recorded in small time intervals; thus, they can be relied upon to amend the roughly recorded pumping/injection rates. M108, M121, M122, M67, and M12A had increased heads in the early stage, probably due to opening of R3, R4 and R5. The Noordbergum effect is also a possible reason for this behavior. These monitoring wells had the strongest responses in the shutdown events, and these particular events may provide information about the hydraulic connectivity between them and R3, R4, R5, and E6. This event is helpful to deconstruct the effects from different pumping and injection wells. It should be noted that the head change plots can reflect that the operations in different wells (e.g., R3 and R11) may have time differences of 0.1 day, which is not recorded in the pumping/injection rates plot. However, to utilize the data from this event in HT inversion, the uncertainty of the source/sink must be considered.

On July 16 2014, the pumping rate in E13 decreased from 300 mpg to 100 mpg. In addition, the pumping rates in E4 and E8 had slight increases. Due to these changes, increased heads were observed in some wells (Figure 5-13(e)), such as M105, M101, M104, M102, and R3M (in descending order in terms of responses). The magnitude of changes can reach up to 3 feet.

Some wells (e.g., M118, M77, and M81) had moderate responses (about 1 foot). The other wells (e.g., M119, R4, R5) had small increases or even decreases of head. It is interesting to see that although M119 and some other wells had large responses in shutdown events, they are not sensitive to the rate change in E13. Thus, this event can provide clear hydraulic connectivity information between E13 and observation wells. This event is probably the most useful in our HT analysis, since the uncertainty in the change of stress intensities (pumping/injection rates) is the least among all events.

On November 24 2014, E23 was closed and the pumping rates at E1 and E13 were slightly increased; also, all the recharge wells slightly reduced their recharge rates. From Figure 5-13(g), we can infer that E23 was closed slightly earlier (0.01 d) and the other changes were passive. This event, as well as that in Figure 5-13(f), confirm that M122 and R11M are connected to well E23. This observation is different from that in the events related to E13, meaning that this event can provide non-redundant information.

On September 24 2014, during the system resume process, well E23 had fluctuations (Figure 5-13(g)). Thus, at the early stage (t<1 d), the head also fluctuated. R3M also had increased head, and while this well is far away from E23, we speculate that the injection rate in R3M was changing, but did not shown in the record. Thus, the data with weak signal or unclear stress should be used with caution.

Another event happened on May 20, 2015, when E6 was shut down and E24 was opened (See Figure 5-13(h)). Similarly, at December 16th 2014, E24 was shut down, and the recharge rates of some wells were reduced accordingly (See Figure 5-11 and Figure 5-13(i)). Although the head changes were small, this still can be used to infer the relative ranking of the responsive signals. For instance, M12A, which is close to E24, was influenced by the shutdown of E24 and had significant decrease of head. However, the data should be de-noised before using in HT inversion.

We selected four events which are reliable and non-redundant (i.e., system resume/recovery; E13 rate reduce; E23 shutdown and E4, E8 slight increase, see Figure 5-13) in the HT inversion.

Figure 5-12 presents the individual variations of head during the monitoring period (June 2014 to July 2015). Throughout this long period, the responsiveness of each observation well is clearly demonstrated. Also, the standard derivation value can quantify the degree of responsive. The responses can be assorted into three categories: Positive responses; negative responses and lack of responses. Positive responses mean the wells are more influenced by pumping wells. Thus, during system shutdown, the head first increases, then decreases. For instance, the wells M108, M110, M119, M121, M122 and M67 have strong positive responses. These wells (except M67) are close to E1, and E2. In terms of vertical intervals, the intervals of these observation wells are below the Pinchout Aquitard (Figure 5-5). That means the effective intervals of the pumping wells may lay in the UZLU confining Unit (Figure 5-5). The typical wells with negative responses are R3M, R4M, R5M, R11M and M6. These wells are either served as recharge wells or close to them. The third type of wells did not respond to the events of stress changes. These wells either have very shallow and short intervals (e.g., P2, P3, P4, P6, P7, P10, S33, B1, and B3, in the UZUU) or very deep intervals (EL4M and M12B, in the lower zone). It is interesting to note that although these shallow or deep wells did not respond to the pumping and injection, the head observations had a long-term trend of increasing. The possible reasons for this may be recharge and regional groundwater flow, but we currently are not able to identify them.

Inversion Model Design

A conventional MODFLOW model was created and calibrated for comparison with the HT model. The domain has 60 columns and 80 rows, with 50m-long quadratic cells. The side boundaries are general head boundaries. A plan view of the domain is depicted in Figure 5-14. The model consists of 21 layers, which have a higher resolution within the depth well screens and lower resolution otherwise. The total depth is 122m.

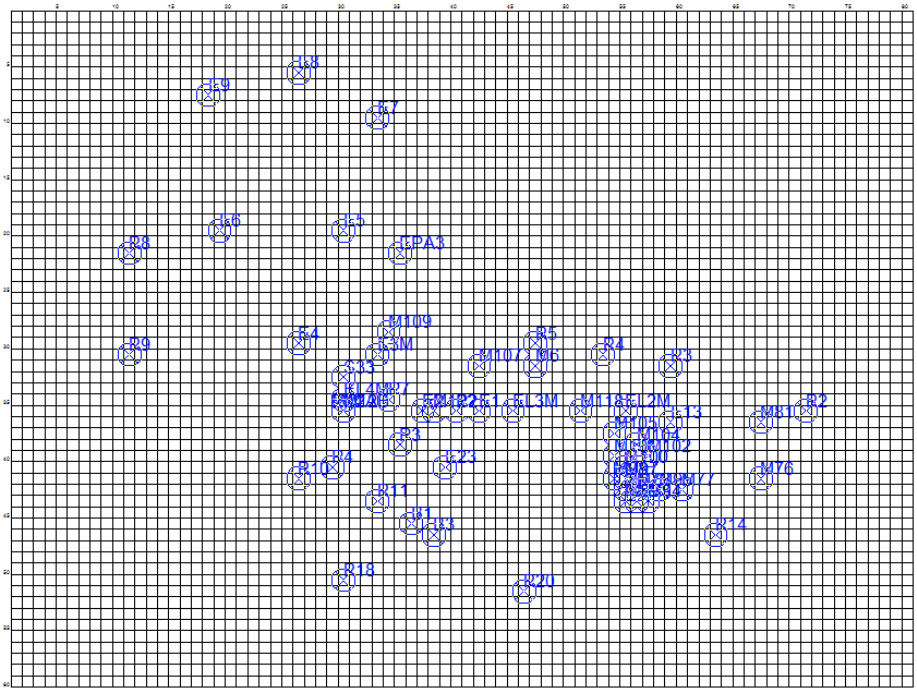


Figure 5-14. Plan View of AFP44 MODFLOW Model

Our HT model is vertically discretized into 20 layers. The 4,000m x 3,000m domain is discretized into 50m x 50m elements. To account for the borehole effects, we adopt the 1-D finite element superposition approach of Sudicky et al. (1995) to our control volume finite element model used in HT inversion.

Cross-correlation analysis

Cross-correlation analysis was utilized to illustrate the information inferred from each piece of observation data. We advocate the use of cross-correlation rather than sensitivity for evaluating the information content in each observation data. As discussed in Mao et al. (2013b) and Sun et al. (2013), the cross-correlation analysis is the sensitivity analysis casted in stochastic framework. It includes not only the sensitivity but also the spatial correlation of parameters to describe the information about the heterogeneity based on head observations. The cross-covariance ($\boldsymbol{\varepsilon}_{dy}(\mathbf{x}_0, \mathbf{x}_j)$) between observation at $\mathbf{x}_0$ and the parameter at $\mathbf{x}_j$ is the summation of the Jacobian ($\mathbf{J}_{dy}(\mathbf{x}_0, \mathbf{x}_k)$) weighted by the covariance of parameters (($\boldsymbol{\varepsilon}_{yy}(\mathbf{x}_k, \mathbf{x}_j)$)). The cross-correction then is obtained via normalizing cross-covariance:

$$\boldsymbol{\rho}_{df} = \left[\operatorname{diag}\left(\boldsymbol{\varepsilon}_{dd}\right)\right]^{-1/2} \boldsymbol{\varepsilon}_{df} \left[\operatorname{diag}\left(\boldsymbol{\varepsilon}_{ff}\right)\right]^{-1/2} \tag{5-1}$$

The cross-correlation map contains information about K heterogeneity conveyed by a head measurement induced by a pumping test. To illustrate the influence of prior information (mean K and covariance of K) on the cross-correlation map, we analyze the cross-correlation maps of AFP-44 site in a two-dimensional case. The correlation patterns are related to the flow conditions (pumping and injection rates) and the observed location. The mean K value is 24 m/d (or geometric mean 3 m/d), mean S_s value is 0.00015 m^{-1} (Zhang and Brusseau, 1999), aquifer thickness is 60m. As discussed in previous section, early time (with respect to the beginning of each event) is hard to locate due to the imprecise record of pumping and injection rates. As a result, it may be very difficult to identify the storage coefficient, as concluded by Sun et al. (2013). In this cross-correlation analysis, we focus on the cross-correlation between lnK at different locations and late-time head observations at different wells. Two scenarios are considered: the first one uses the initial pumping and injection strategies (June 2014), and scenario 2 simulates the decrease of pumping rate in E13 during July 2014.

We assume the isotropic correlation scale of K is 300 m. Contour maps of the cross-correlation between the head at observation wells B1, E3M, EPA3, M6, M95, P3 and lnK perturbation everywhere in the domain are plotted in Figure 5-15. These observation wells are selected so that they can represent observations at different regions. Figure 5-15(a) plots the cross-correlation between lnK and late-time head observation at B1, which is located at the south region of AFP44. Observation well B1 mainly receives recharges from R10, R11, R18, and R20 (all in southern of B1), and supply water to E1, E2, E23, and E24 (all in northern of B1). Apparently, if there is a low K barrier between the pumping wells and B1, then B1 tends to have high head since it is not influenced by the pumping wells. Thus, the cross-correlation map has negative values between B1 and surrounding extraction wells. It is also observed that the head in B1 has positive correlation with lnK at the region between E13 and R2, and the region between E4 and R9. The reasoning is this: if these pairs of pumping and injection wells have well hydraulic connection, then B1 does not supply water to those nearby extraction wells. Figure 5-15(b) plots the cross-correlation between lnK and late-time head observation at E3M. The pattern is different because E3M is outside of the major pumping-injection dipoles (E1, E2, E4, E23, E24 and R11, R10, R18 and R20). Since E3M is closer to these pumping wells, the cross-correction between head in E3M and lnK at the region of the major pumping-injection dipoles is positive, meaning that water is directly supplied by injection wells, rather than E3M. By the same token, the cross-correlation patterns of observations in EPA3, M6, M95 and P3 can be explained. Overall, the general correlation patterns are very similar in E3M, EPA3, M6, M95 and P3, since the gradient is relatively flat when there are multiple sources and sinks.

Figure 5-16 investigates the situation when E-13 has reduced its pumping rate. In Scenario 2, most of the cross-correlation maps do not change, except M95 (Figure 5-16(e)). In this situation, the cross-correction between head at M95 and the lnK at its western region is negative. The reason is that this time, both R2 and E13 are designed with large injection and pumping rates (both wells are responsible for 20% of the total pumping and injection rates). When E13 reduces its pumping rate by 66%, the excess water from R2 will recharge E1, as indicated by the streamlines. If the lnK between R2 and E13 is higher, and the lnK between E13 and E1 is smaller, the observation M95 tends to have high head due to the recharge of R2 and isolation from E1. This means that change of pumping rate will partially change the flow field, and provide non-redundant information.

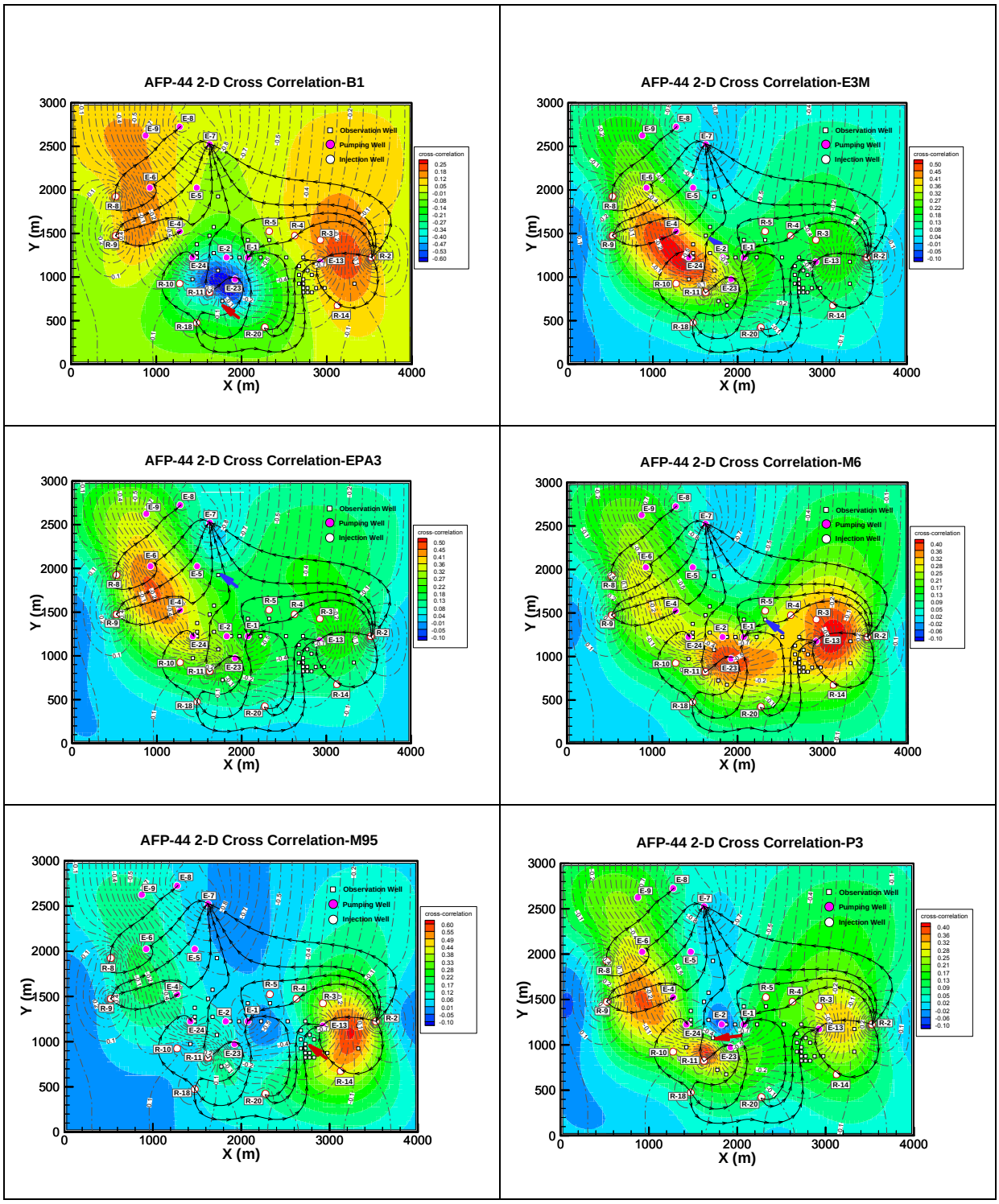


Figure 5-15. Cross-correction Analysis in Scenario 1

Arrowed blacked lines are streamlines. Grey dashed lines are equipotential lines of head. The contour indicates the cross-correction between lnK and observed head at different observation locations.

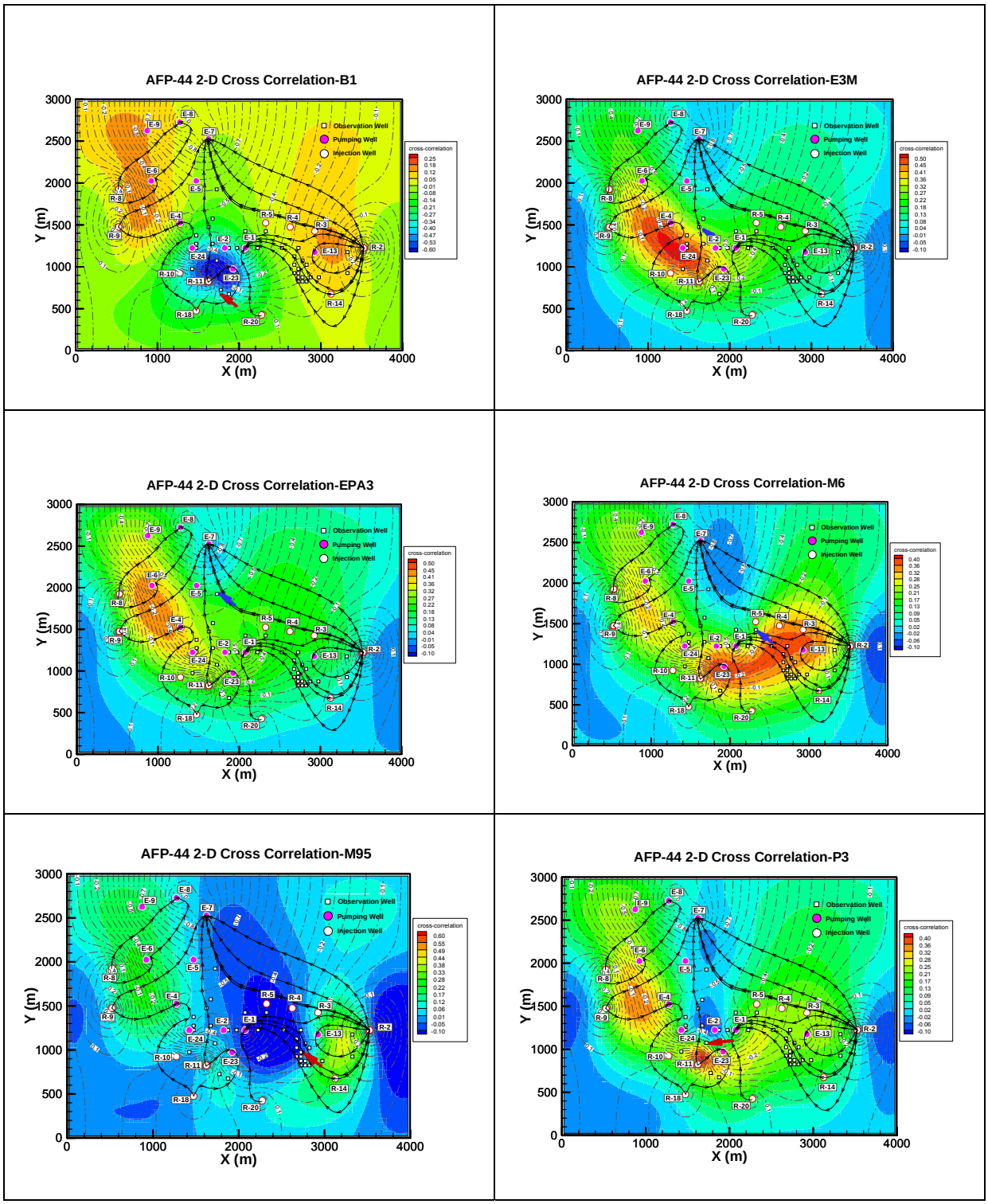


Figure 5-16. Cross-correction Analysis in Scenario 2

Arrowed blacked lines are streamlines. Grey dashed lines are equipotential lines of the head. The contour indicates the cross-correlation between lnK and observed head at different observation locations.

Parameter Estimation by Model Inversion

We first built a simplified 2-D model with an aquifer thickness of 60m as the first step to investigate the influence of well settings and initial prior information. Figure 5-17 shows the estimated K- and Ss-fields. Figure 5-18 shows the estimated residual variance of the K- and S_s-fields. Figure 5-19 shows a comparison of the simulation head and monitored responses to the pumping tests used for calibration.

Some observed features in the data can be reasonably explained. A high K zone connects E6, E4, E24, R10 and R11. Thus, due to starting E24, a significant decrease of the head (>0.1 m) was seen in M12A, M122, and R11M. E13 was located in a low K zone. The reason is that when the rate of E13 decreased by 66%, some surrounding observations (i.e., M101 and R3M) had a large increase of head (>0.5 m). The strong horizontal heterogeneity may be the artifacts of using a 2-D model.

After assimilation of HT data, the uncertainty of lnK seems to be reduced between pumping/injection and observation wells. The residual variance values are less than 0.2 at some locations where the observation wells are clustered or in locations near the pumping or injection wells. This means that at these locations, the uncertainties are reduced by 80% using HT data. On the other hand, the uncertainty in the outer region with wells turned on maintains the same value as the unconditional uncertainty (1.0). This means that the HT data have no constraints on this region. In contrast, the reduction of the $\ln S_s$ variance is limited to the region with observation wells. This is consistent with the cross-correlation analysis in Sun et al. (2013).

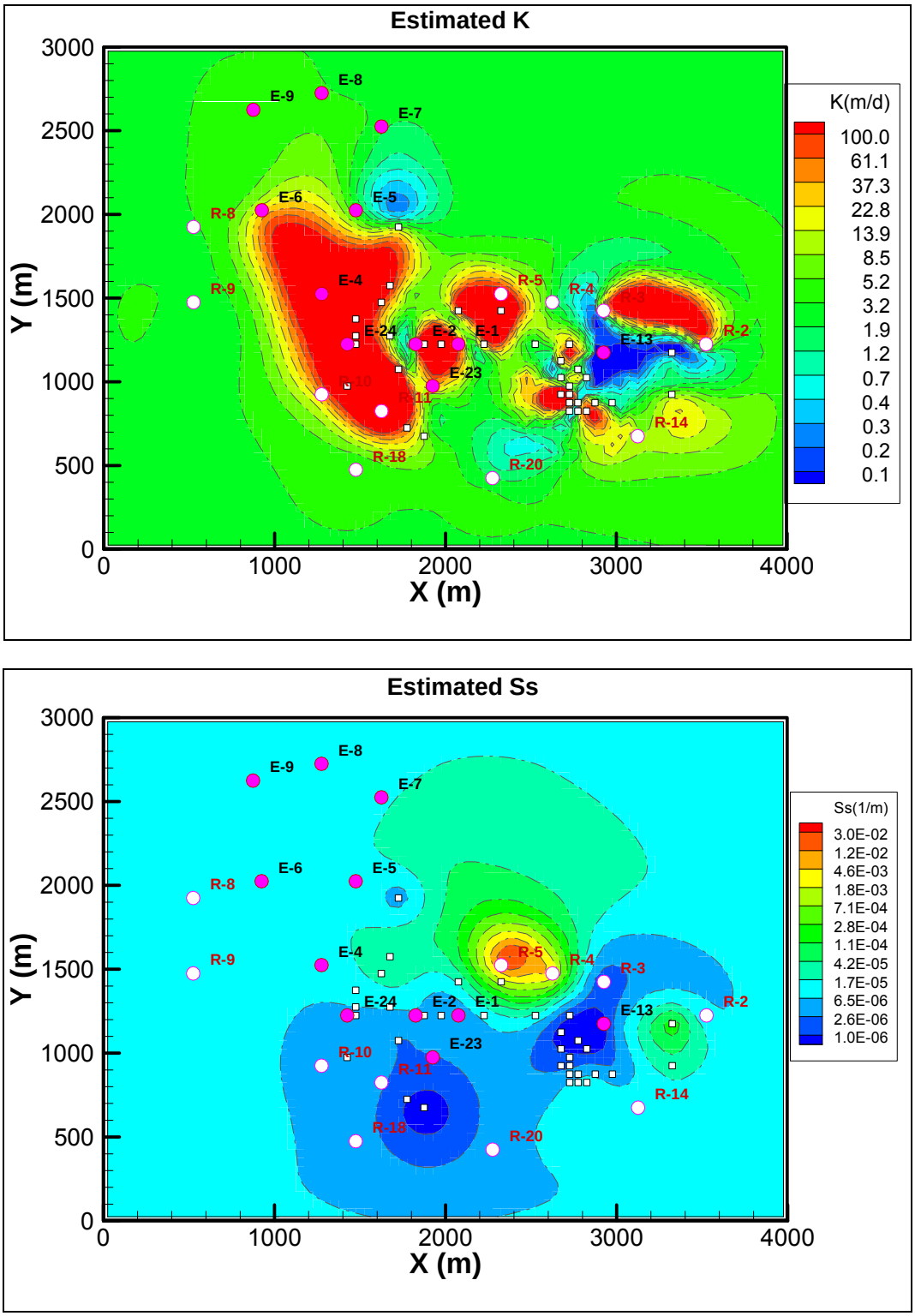


Figure 5-17. Estimated K and S_s Fields Used in the 2-D Case with Four Events Using Collected Field Data

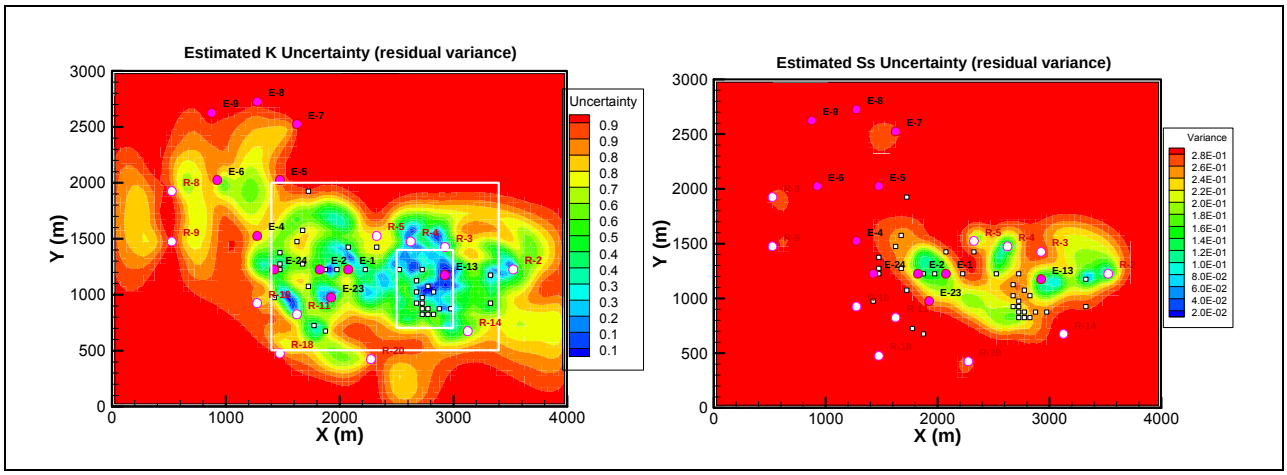


Figure 5-18. The Uncertainty (residual variance) of Estimated lnK and lnS_s for 2-D Inversion Case with Field Data

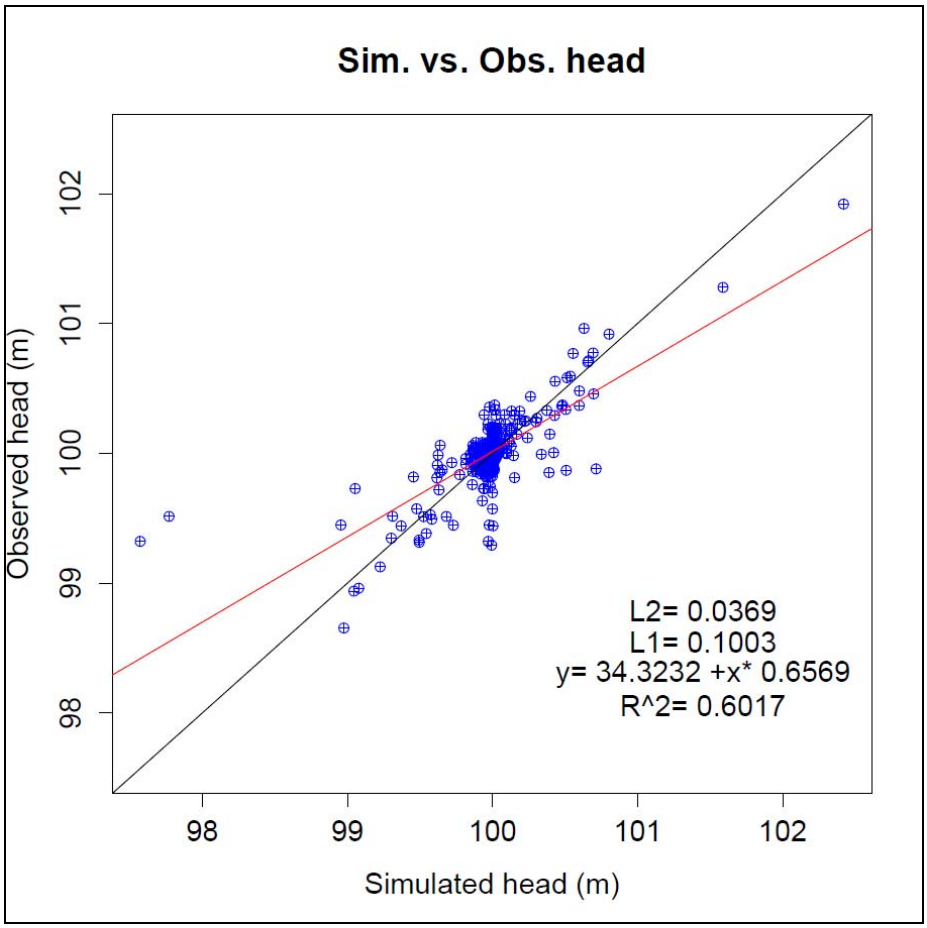


Figure 5-19. Calibration Map of Head Scatterplot in the Real case with Four Events

After the 2-D HT analysis, we performed three-dimensional HT inversion using the same datasets as those in the 2-D case. The estimated K still exhibits a low K zone (Figures 5-20 and 5-21) in the southeast region, which is consistent with the 2-D results. However, the results also show strong vertical heterogeneity, which is not shown in 2-D case. It should be noted that the figures showing the estimates have different scales for vertical and horizontal directions.

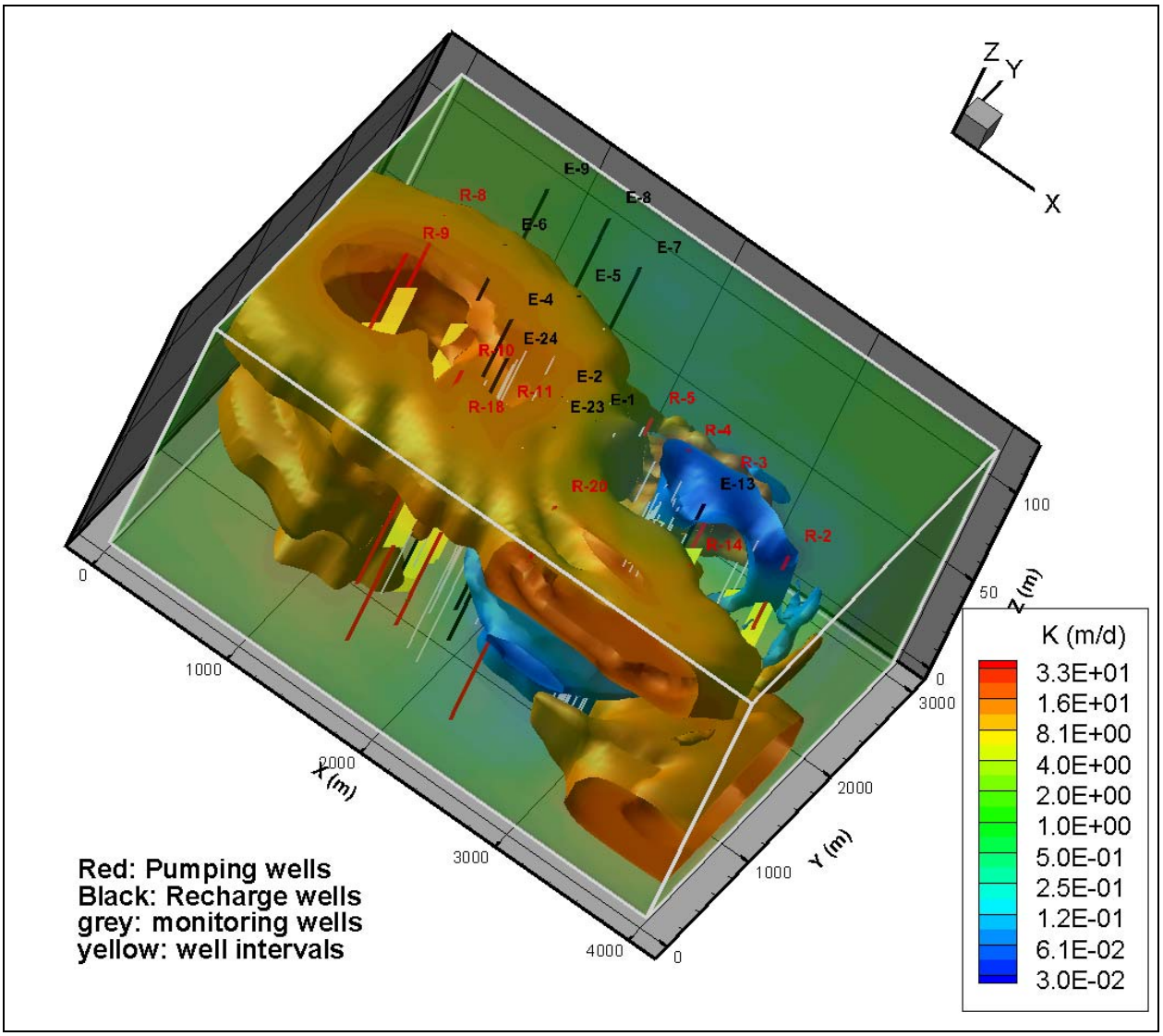


Figure 5-20. Isosurfaces of the Estimated High K and Low K Zones Using 3-D Model with Consideration of the Long Borehole Interval

The color values are lnK perturbations (difference compared to the mean value ln(3))

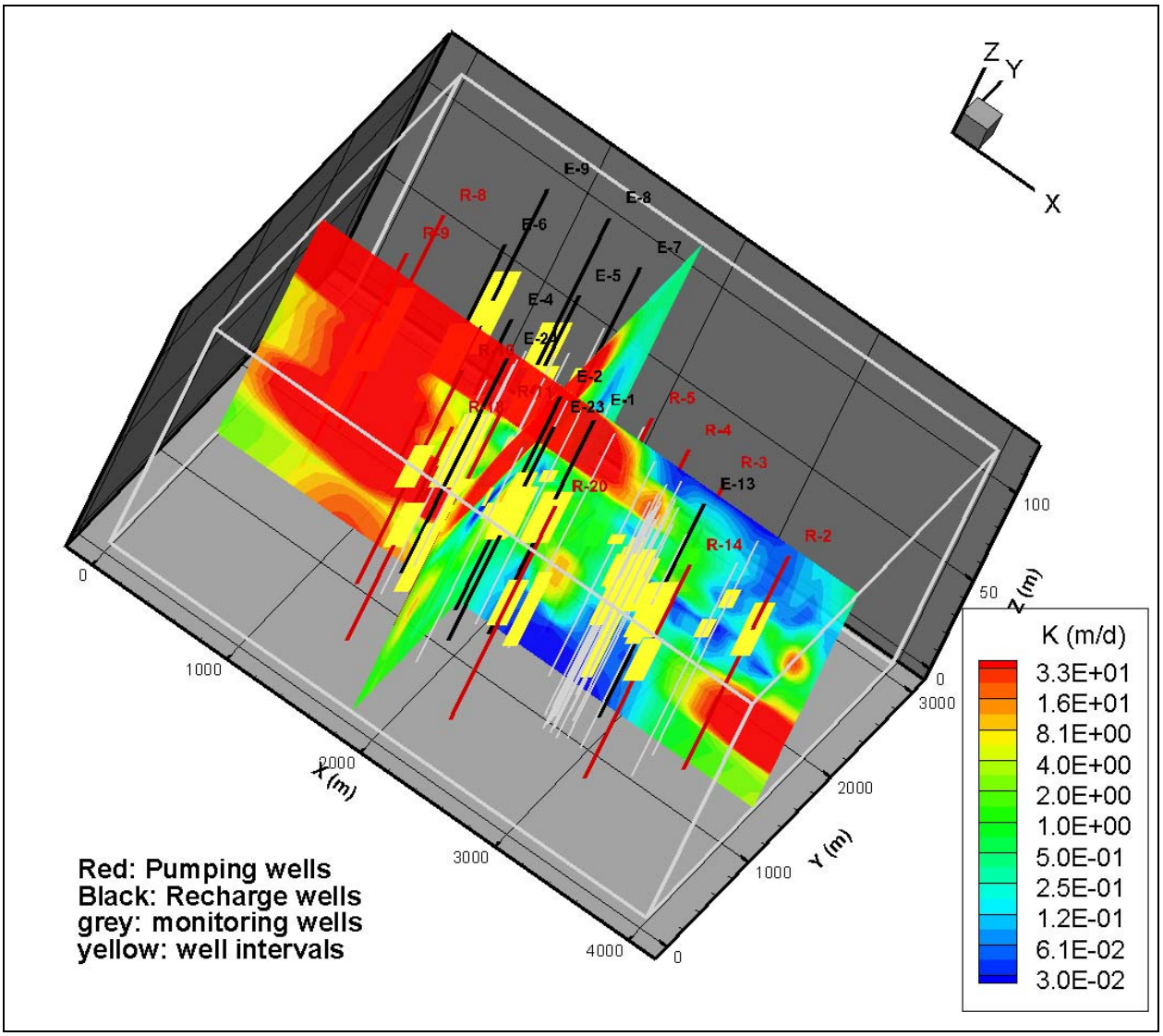


Figure 5-21. Slices of the Estimated High K and Low K Zones Using 3-D Model with Consideration of the Long Borehole Interval

The estimated S_s pattern is difficult to relate to geological information (Figure 5-22). Illman et al. (2009) showed that the estimated K and S_s for a fractured medium have strong negative correlation. However, in this case with porous medium the correlation seems to be insignificant.

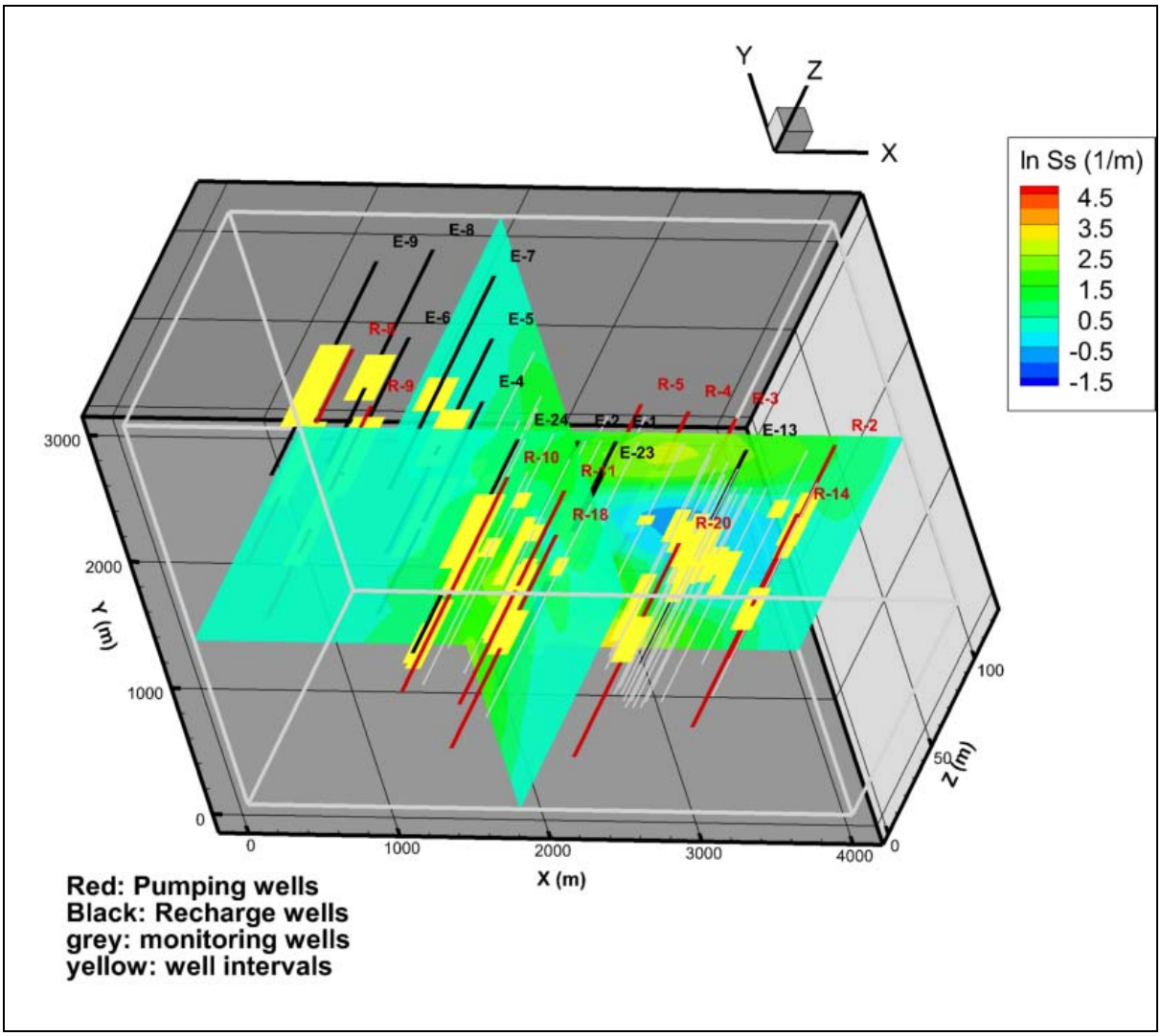


Figure 5-22. Slices of the Estimated S_s

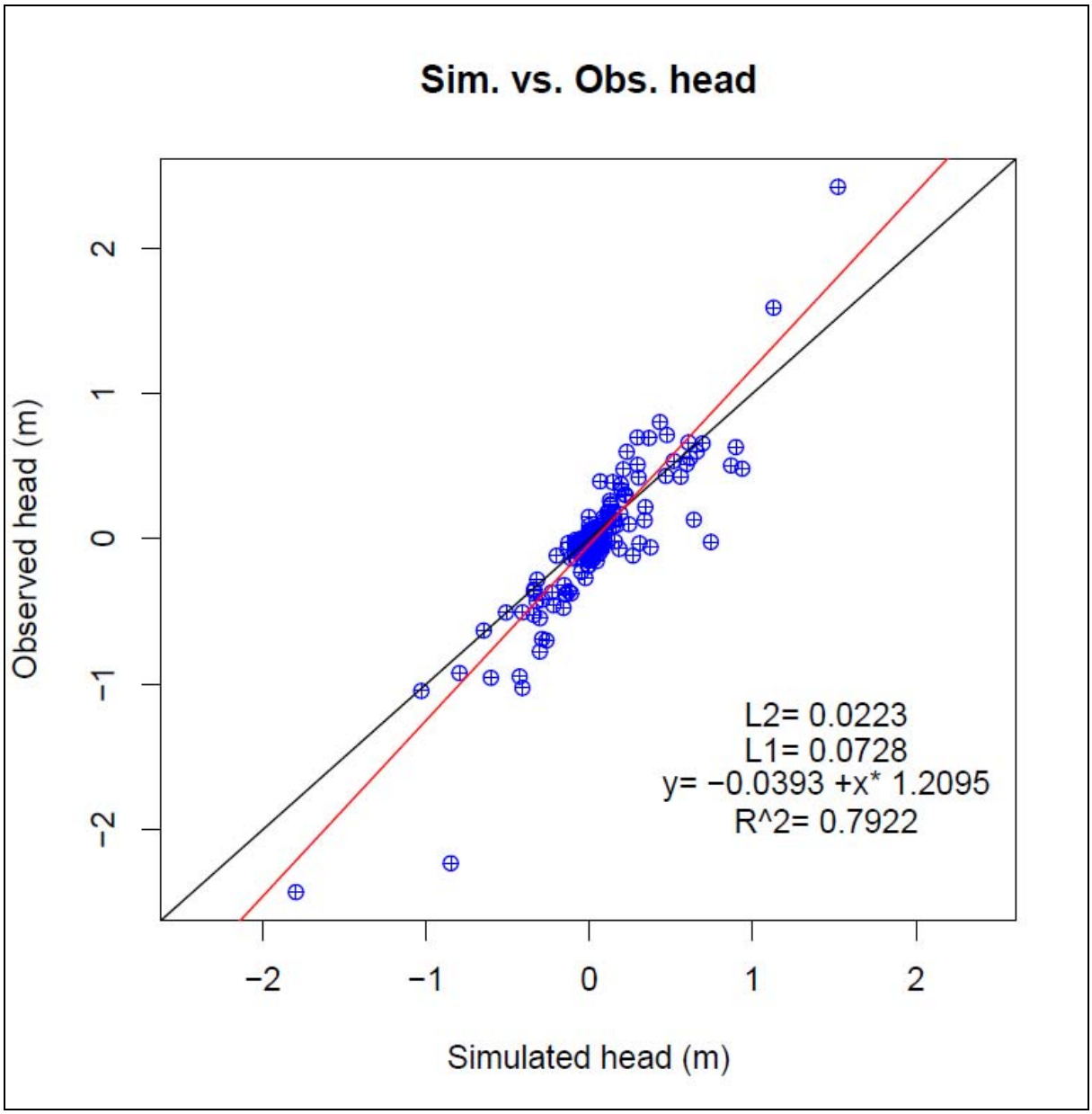


Figure 5-23. Calibration Map of Head Scatterplot in the Real Case with Four Events

Figure 5-23 shows the comparison of the model-predicted head responses to the recorded responses to the pumping tests. In comparison to the 2-D results shown in Figure 5-18, 3-D HT results are noticeably better than those in the 2-D case. This indicates that the 3-D conceptual model is more realistic for the characterization of the flow and heterogeneity at this site.

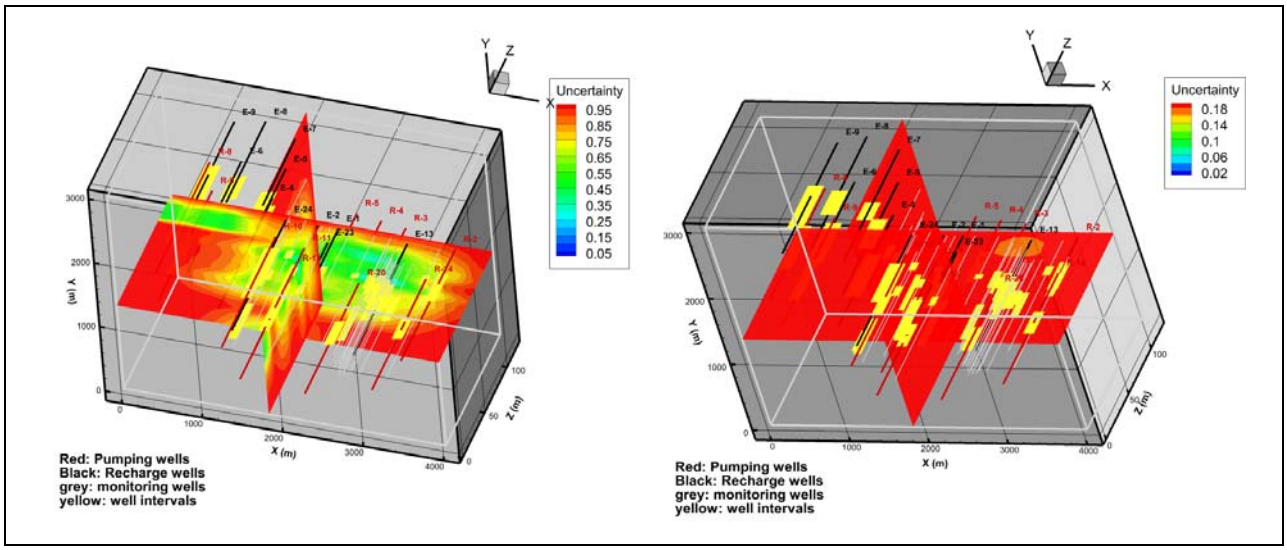


Figure 5-24. The uncertainty (residual variance) of estimated lnK (left) and lnS_s (right) for 3-D inversion case with field data. The original variances for lnK and lnS_s are 1.0 and 0.2

Figure 5-24 displays the uncertainty (measured by residual variance) of estimated lnK and lnS_s in the 3-D inversion with field data. The original variances for lnK and lnS_s are 1.0 and 0.2. The estimated uncertainty maps are similar to those in 2-D cases. The estimated lnK residual variance values are reduced at the well locations to a significant extent. In contrast, the estimated lnS_s residual variance values only decrease at the locations of the observation wells.

Page Intentionally Left Blank

6.0 PERFORMANCE ASSESSMENT

The HT performance is evaluated according to the performance objectives defined in section 3.

6.1 PERFORMANCE OBJECTIVE: DEMONSTRATE HIGHER ACCURACY OF HT AGAINST CONVENTIONAL SITE CHARACTERIZATION TECHNIQUES

The effectiveness of site characterization methods is measured by the extent to which the observed pumping test responses match the predictions based on the estimated hydrogeologic parameters. The results from the demonstrations at both the AFP44 and NCRS sites demonstration confirmed that the HT is more accurate than conventional site characterization techniques.

6.1.1 NCRS at UW, Canada

We used the estimated K-fields from all cases to predict drawdowns of the seven pumping tests (PW3-1, PW 3-4, PW5-5, PW1-3, PW1-5 and PW1-7) not used in model inversion. These tests were selected from pumping ports located in different areas of the modeling domain. The scatterplots of observed and simulated drawdowns are shown in Figure 6-1. Linear fit results and L_1 and L_2 norms are also included to evaluate the overall prediction performance for the seven pumping tests. A perfect prediction would show the simulated drawdowns on the 45-degree line, which is achieved by HT to a larger extent than by the conventional methods.

Examinations of Figures 6-1(a) to 6-1(h) revealed that the results of drawdown predictions improve gradually from the effective parameter approach (Case 1) to the highly parameterized approach based on HT inversion (Case 3). The isotropic effective parameter model Case 1a yielded results that have the highest L_1 and L_2 norms for drawdown prediction while considering anisotropy in Case 1b or using a 5-layer geological model improved the results. Additionally, prediction results of the complex 19-layer geological model ranked in the middle. The geostatistical model has the lowest L_1 and L_2 norms with very close prediction performances. Specifically, geostatistical inversion Case 3d, using the uncalibrated geological model populated with permeameter K data as a prior distribution, performed the best. When geologically distributed K values were used as prior distributions, it is interesting to note that the geostatistical inversion Cases 3b, 3c and 3d performed quite closely in terms of model calibration and validation and only slightly better in terms of R^2, L_1 and L_2 norms than Case 3a, in which a uniform K prior mean value was used, given the differences in the estimated K tomograms (Figures 5-8c-f).

The results clearly indicate that the HT predictions outperform those of conventional models. The HT predictions are unbiased and have smaller Root-mean-square of drawdown prediction errors. These results are consistent with findings in the laboratory sandbox study of Illman et al. (2015) that the geostatistical inversion approach performed the best in terms of drawdown predictions when compared with effective parameter and geological modeling approaches.

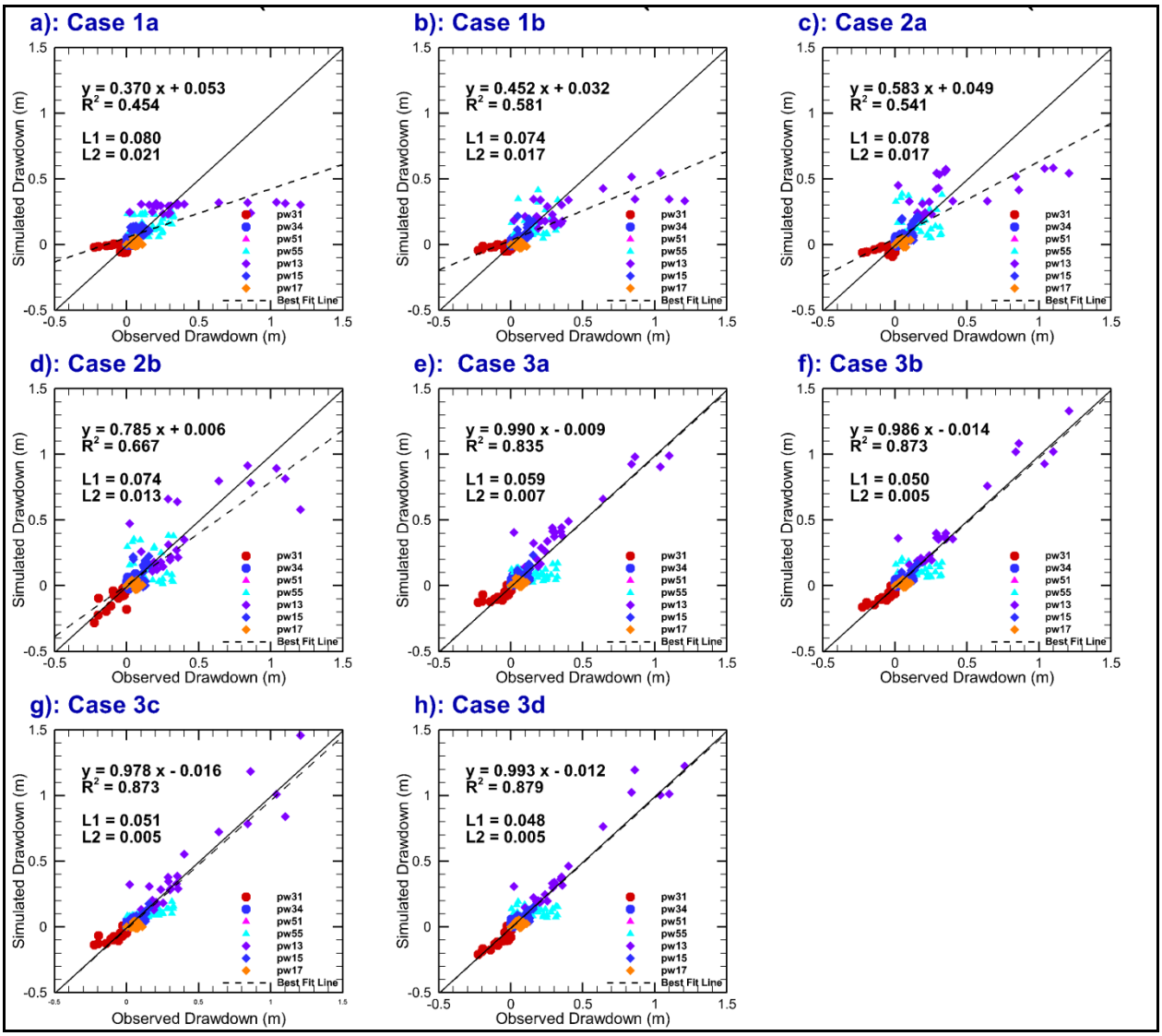
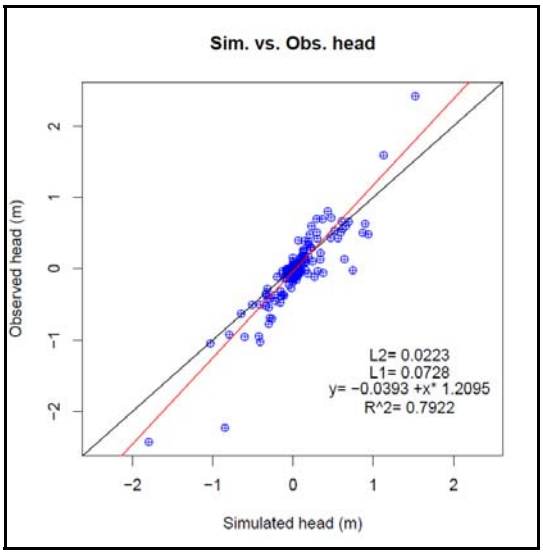


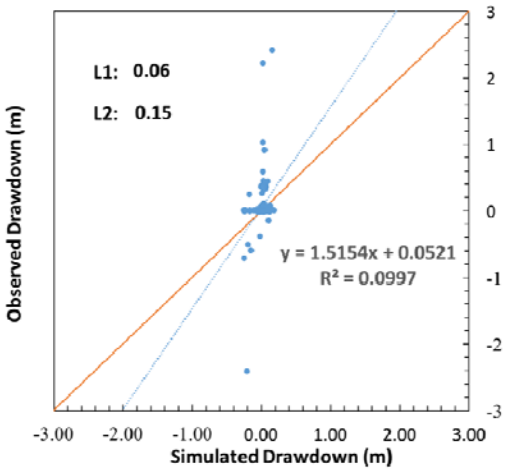
Figure 6-1. Scatterplots of Observed Versus Simulated Drawdowns for Model Validation Using Seven Pumping Tests ¶or the: (a) Isotropic Effective Parameter Model; (b) Anisotropic Effective Parameter Model; (c) Geological Model with Five Layers; (d) Geological Model with 19 Layers; (e) SimSLE Starting with K = 8.0 × 10^{-6} m/s as Prior Mean; (f) SimSLE Using the Calibrated Five-layer Geological Model as Prior Distribution; (g) SimSLE Using the Calibrated 19-layer Geological Model as Prior Distribution; and (h) SimSLE Using the Uncalibrated 19-layer Geological Model Assigned with Permeameter K Values as Prior Distribution.

6.1.2 AFP44 Site in Tucson, Arizona

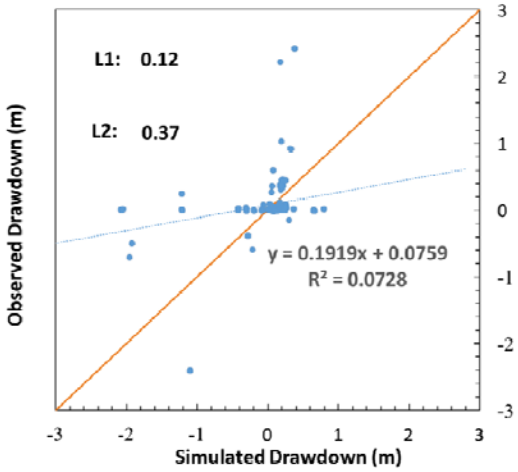
Figure 6-2 shows a similar plot for the AFP44 validation test. The HT model results are shown in Figure 6-2(a). The homogeneous and layer model results are shown in Figure 6-2(b) and (c). The mean square error of the HT result is persistently smaller. The figure clearly shows that the HT predictions outperform the predictions from conventional models.



(a) HT model



(b) Homogeneous model



(c) Layered model

Figure 6-2. Simulated Versus Observed Drawdown of AFP44 for Validation Pump Test

The uncertainty of the drawdown predictions was estimated with a first-order approximation and was compared to the prediction error. Figure 6-3, Figure 6-4 and Figure 6-5 show the observed and predicted drawdown curves for the AFP44 validation data by observation location for the HT model, the homogeneous model, and the layer model result, respectively. Each plot also shows the upper and lower bounds of the prediction standard variation. It is evident that, for the HT result, the difference between prediction and observation is within the standard variation bounds except for a few locations, whereas the conventional methods fail to meet this requirement at most of the observation locations.

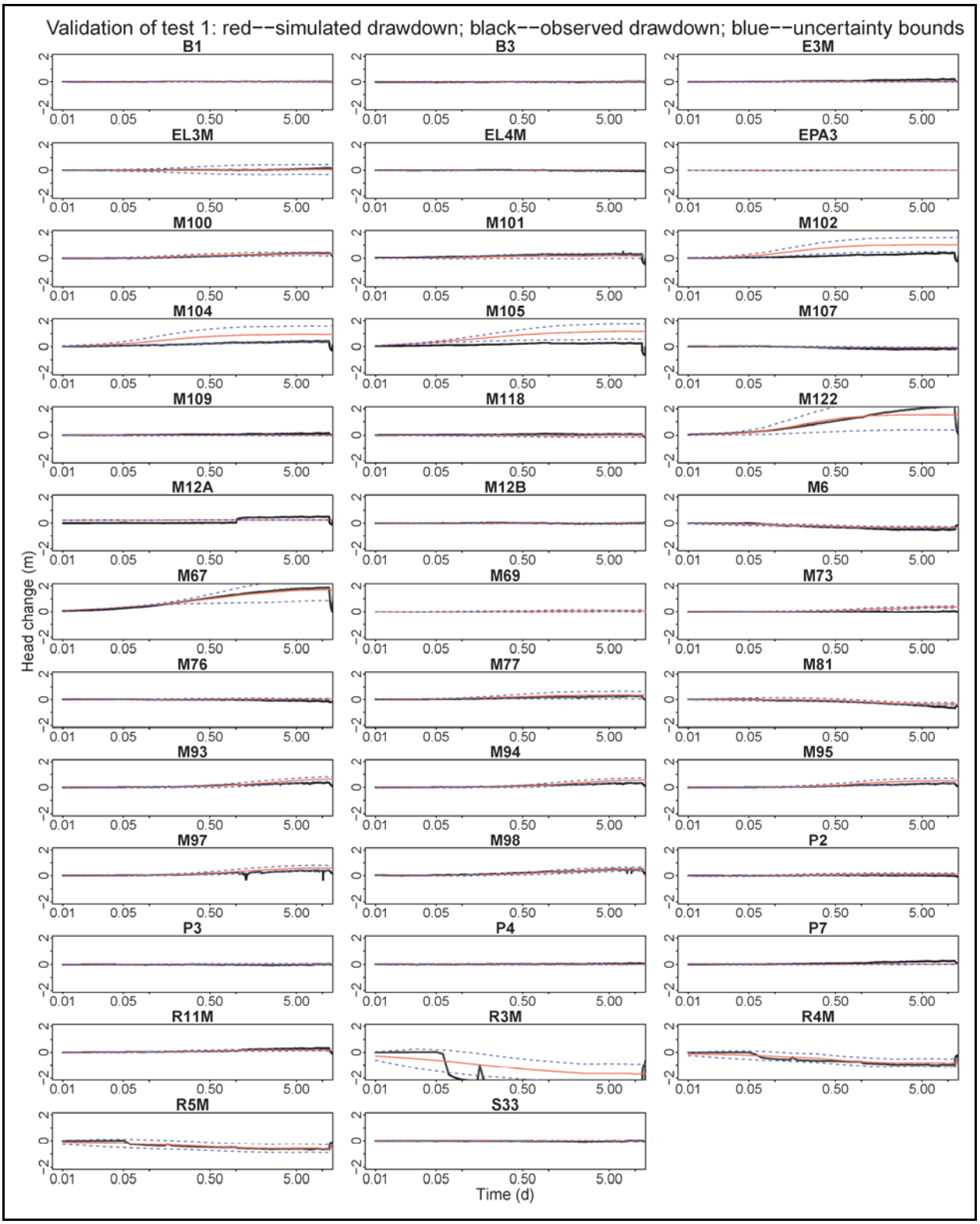


Figure 6-3. Predicted Drawdown, Observed Drawdown, and Prediction Standard Deviation Bounds per Pump Location of the AFP44 Validation Test for the HT Result

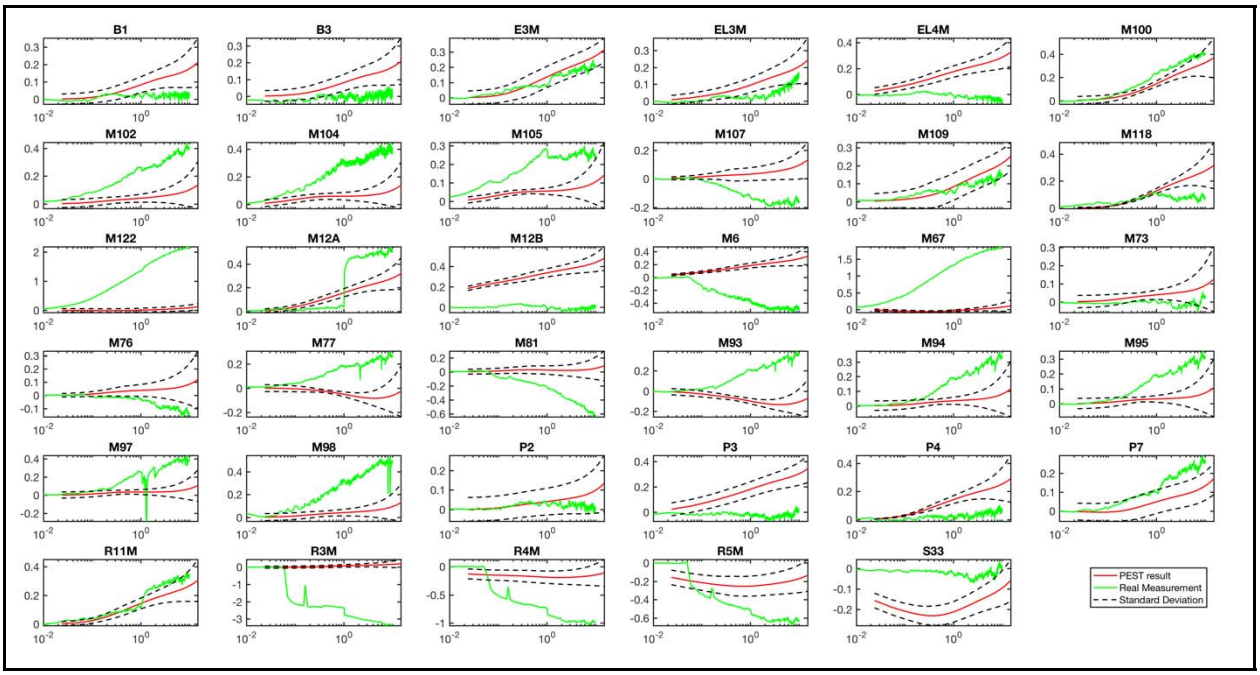


Figure 6-4. Predicted Drawdown, Observed Drawdown and Prediction Standard Deviation Bounds per Pump Location of the AFP44 Validation Test for the Homogeneous Model Result

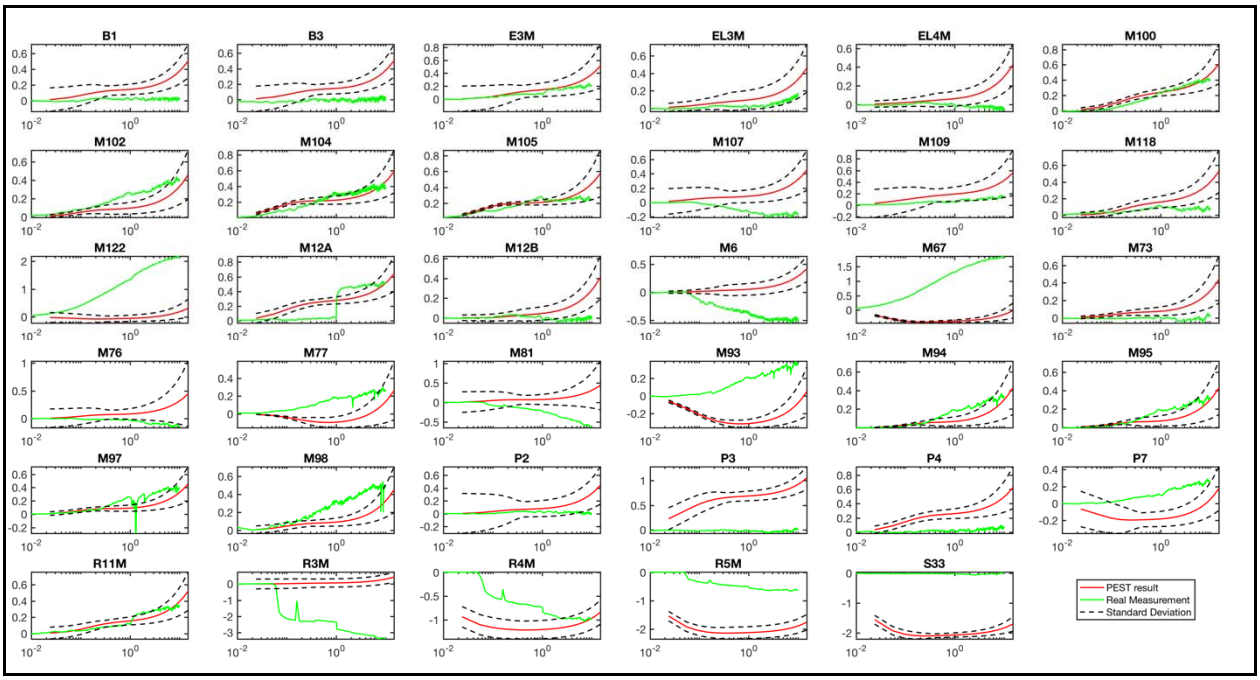


Figure 6-5. Predicted Drawdown, Observed Drawdown and Prediction Standard Deviation Bounds per Pump Location of the AFP44 Validation Test for the Layer Model Result

6.2 PERFORMANCE OBJECTIVE: DEMONSTRATE LOWER UNCERTAINTY OF HT AGAINST CONVENTIONAL SITE CHARACTERIZATION TECHNIQUES

Estimation variance of K- and S_s-values is a measure of the uncertainty associated with estimation methods. The results from the demonstrations at both the AFP44 and NCRS sites confirmed that the HT results are less uncertain than the results from conventional site characterization techniques.

6.2.1 NCRS at UW, Canada

For NCRS, the smallest parameter variance of both the homogeneous and layer model is 1.00 for K in natural scale. Converting the highest single-element variance obtained by HT (see Figure 5-12) into a uniform variance over one layer gives 0.33 for K, which is already lower than the uniform average over a set of layers, let alone over the whole model.

6.2.2 AFP44 Site in Tucson, Arizona

The variance of K and S_s in the homogeneous model are 1.02 for K and 2.13 for S_s in the natural scale. In the layer model, the smallest variances for K and S_s are 1.10 and 1.05, respectively. Figure 5-17 shows the variance of the K- and S_s field estimated by HT. Converting the highest single-element variance into a uniform essemble variance over one layer gives 0.0005 for K and 0.0002 for S_s, which is lower than the smallest variances for K and S_s for the layer model.

6.3 PERFORMANCE OBJECTIVE: ILLUSTRATE CONSISTENCY OF HT RESULTS WITH LITHOLOGIC/GEOLOGIC DATA

The consistency of the K-field delineated by HT with available core information is a qualitative indication of the accuracy of HT. The results from the demonstrations at both the NCRS and AFP44 sites confirmed the consistency of HT results with the current spatial distribution knowledge of the more permeable and less permeable regions.

6.3.1 NCRS at UW, Canada

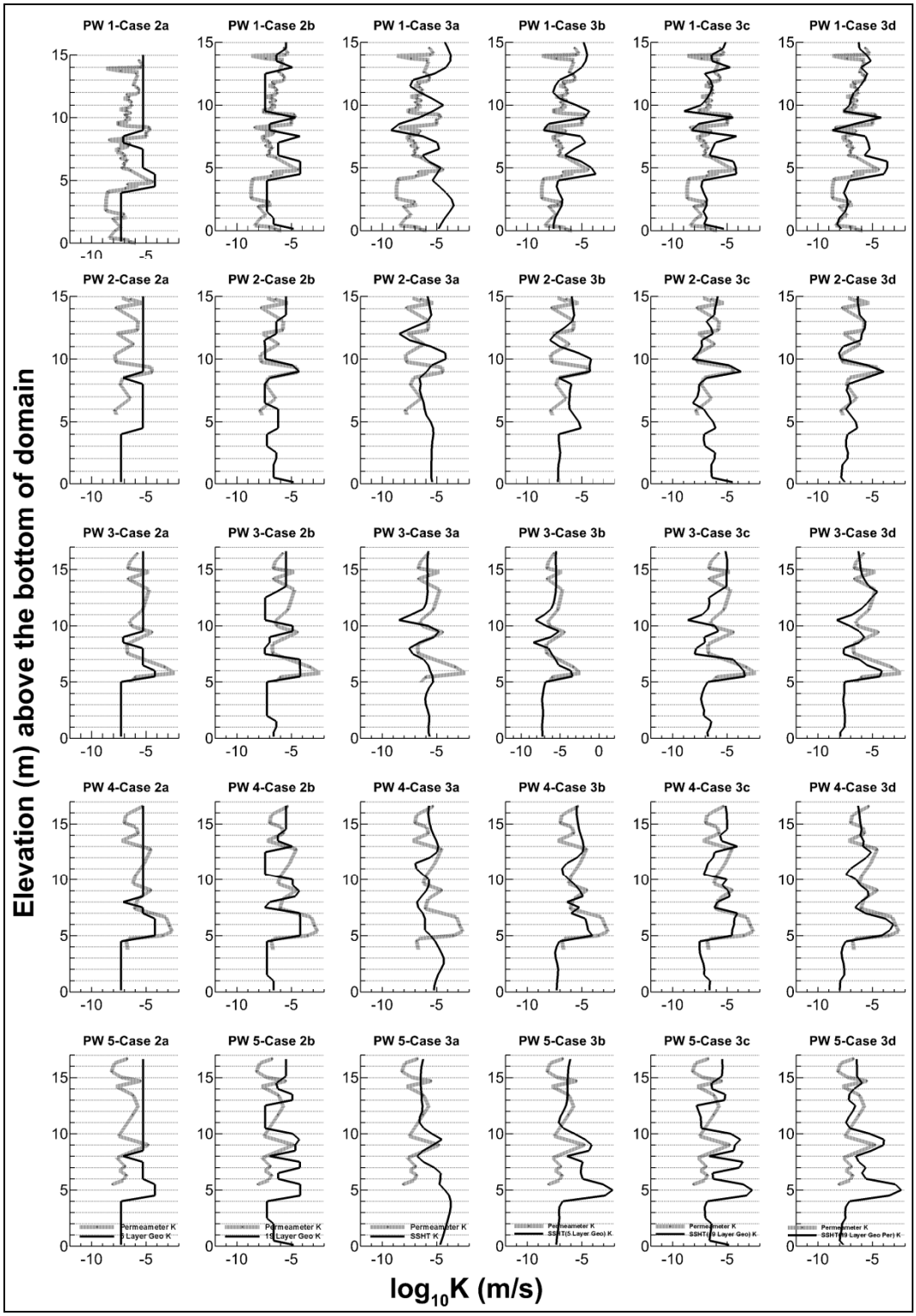
We compared the estimated K values of all scenarios from Cases 2 and 3 to permeameter K values obtained along the CMT and PW wells, as shown in Figure 6-6. This kind of comparison enables examinations of both intra- and inter- layer K variations among different subsurface characterization approaches.

A comparison of the results from Case 3a to permeameter test K values reveals that when a homogeneous K-field is used as the prior mean, the geostatistical inversion approach has only captured the general features of high and low permeable layers within the range of 5 m to 12 m, and K estimates for the area away from the well field are relatively smooth. The main reason for this is that no observation data are available to update the K estimates during SimSLE inversion (Xiang et al., 2009). However, as shown in Figure 6-6, when geologically distributed K-fields are used as prior distributions (Case 3b, Case 3c and Case 3d), the fits between the estimated and permeameter tested K values for all CMT and PW wells are consistently improved.

The improvements are most obvious for the high K zone located between 4 m and 7 m above the bottom of the modeling domain, as well as the low K zone near the bottom domain. This result supports the finding by Tso et al. (2016) through a synthetic case study that the prior knowledge of the site-specific geological structures can be important for resolving the correct aquifer features. Additionally, we find from Figure 6-6 that the fit of K profiles in Case 3b with a 5-layer geological model used as a prior distribution in the geostatistical inversion approach is comparable to those from Cases 3c and 3d, in which a 19-layer geological model is used. This finding indicates that a simple geological model reflecting the general geological structure may be sufficient for use as a prior distribution in geostatistical inversion approaches to characterize heterogeneity within the area of interest. Another important feature of the estimated K tomogram of Case 3a is the incorrect mapping of the clay zone at the bottom of the simulation domain; this is the same finding as the previous studies of Berg and Illman (2011b, 2013). This is so, despite the additional steady-state drawdown data from the pumping tests at PW1-6, in which the pumping took place in the bottommost low K zone and was included in inverse modeling. However, we find that the use of transient information (not shown here) correctly resolves the bottom low K zone with pumping test data alone. Overall, the above comparisons suggest that the use of geological data is helpful for the geostatistical inversion approach for HT investigations, in preserving structural features of the hydraulic property field.

By contrast, the K profiles obtained from both calibrated conventional models showed some inconsistency to permeameter-tested results along nine wells. Such inconsistency could be attributed, on the one hand, to using geological zonation with each layer as homogeneous, and on the other hand, to the compensation effect of parameter values to structural errors (Refsgaard et al., 2012). These results collectively suggested that calibration of geological models interpolated from borehole logs to multiple pumping tests is useful in terms of providing general K estimates of the field. However, because the stratigraphy of geological models is fixed in this study, fine-scale variability in K within each layer cannot be captured.

A)


PW 1-Case 2a
PW 1-Case 2b
PW 1-Case 3a
PW 1-Case 3b
PW 1-Case 3c
PW 1-Case 3d
PW 2-Case 2a
PW 2-Case 2b
PW 2-Case 3a
PW 2-Case 3b
PW 2-Case 3c
PW 2-Case 3d
PW 3-Case 2a
PW 3-Case 2b
PW 3-Case 3a
PW 3-Case 3b
PW 3-Case 3c
PW 3-Case 3d
PW 4-Case 2a
PW 4-Case 2b
PW 4-Case 3a
PW 4-Case 3b
PW 4-Case 3c
PW 4-Case 3d
PW 5-Case 2a
PW 5-Case 2b
PW 5-Case 3a
PW 5-Case 3b
PW 5-Case 3c
PW 5-Case 3d
Elevation (m) above the bottom of domain
log10K (m/s)
Permeameter K
6 Layer Geo K
15 Layer Geo K
SSHT K
SSHT(5 Layer Geo) K
SSHT(6 Layer Geo) K
SSHT(9 Layer Geo Per) K

B)

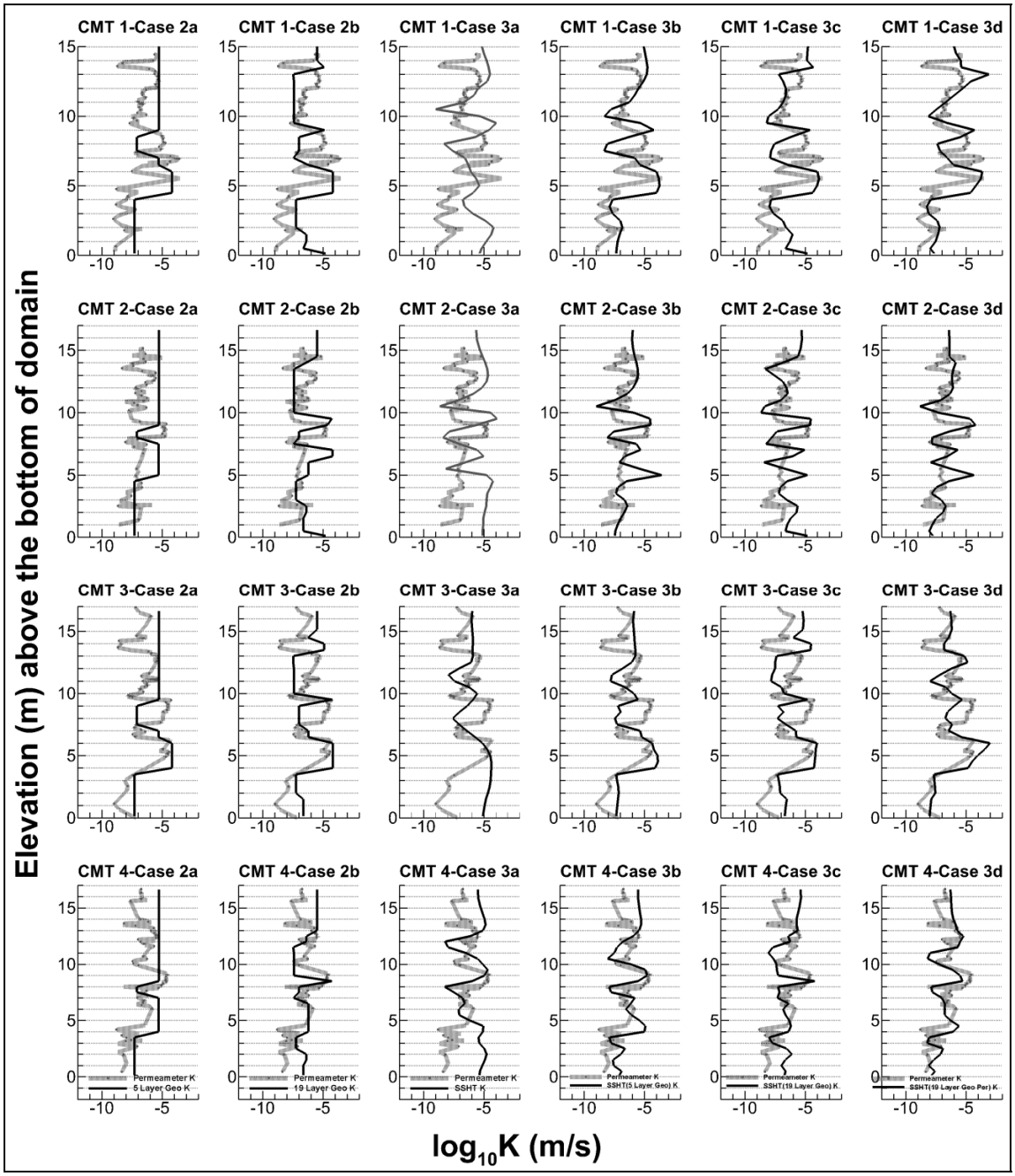


Figure 6-6a,b. Vertical log10K Profiles Along Nine Boreholes of (a) PW Wells and (b) CMT Wells, for Different Calibration Cases

Case 2a: the 5-layer geological model; Case 2b: the 19-layer geological model; Case 3a: SimSLE starting with an uniform K = 8.0 × 10^{-6} m/s; Case 3b: SimSLE using the calibrated 5-layer geological model as prior distribution; Case 3c: SimSLE using the calibrated 19-layer geological model as prior distribution; and Case 3d: SimSLE calibration case using the uncalibrated 19-layer model assigned with permeameter test K values for each layer as prior distribution.

We plotted K estimates for Case 3d models along cross section D-D' (shown in Figure 6-7) for a detailed comparison to the geologic model cross-section. The location of this cross-section is shown in Figure 4-4.

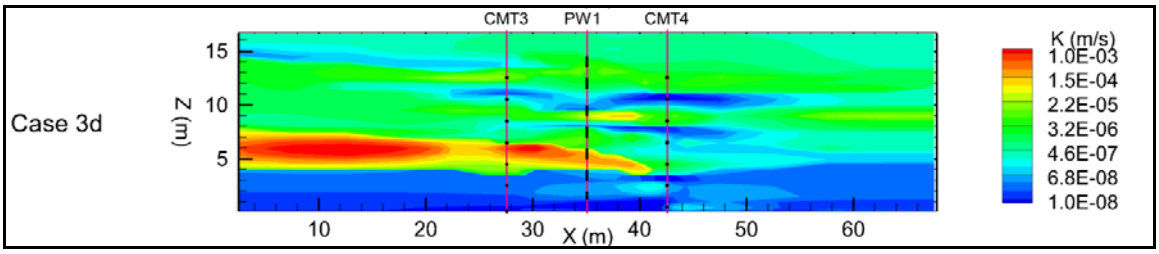


Figure 6-7. Comparison of K Estimates along D-D' Cross Section with Geologic Model

6.3.2 AFP44 Site in Tucson, Arizona

Figure 6-8 shows an overlay map of the spatial distribution of K-values estimated by HT on an aerial photo of the AFP44 site. The results are consistent with the knowledge that more permeable regions (shown in red in the figure) delineated by HT match the regions with more coarse-grained soils and higher well yield. The less permeable regions (shown in blue) are consistent with the regions with more fine-grained soils. The low-K region delineated by HT in the vicinity of M-116 and M-117 is consistent with the area where hydraulic fracturing was performed in 2015-2016 to enhance the recovery of chemicals in the fine-grained soils.

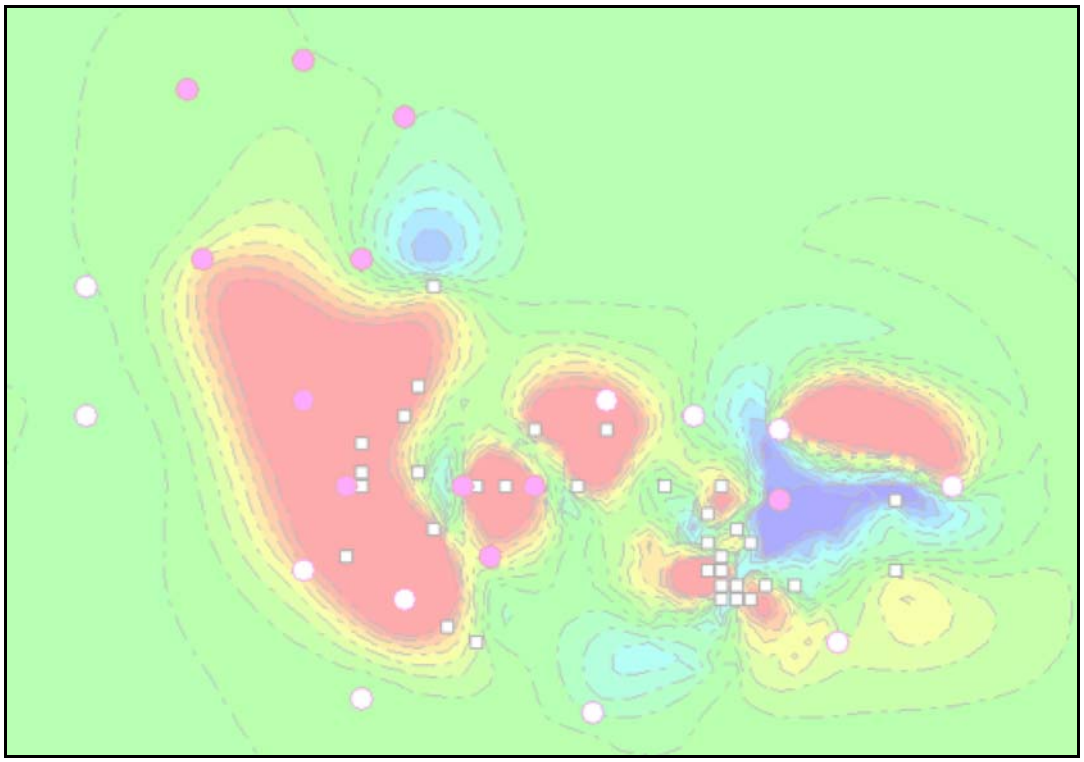


Figure 6-8. Spatial Distribution of K-values Delineated by HT at AFP44 Site

Figure 6-9 shows the delineated K-field in different resolutions. It shows that HT produces results in a resolution consistent with the spacings of the wells in the HT survey. If the wells are spaced closer, HT delineates the K-field in higher resolution.

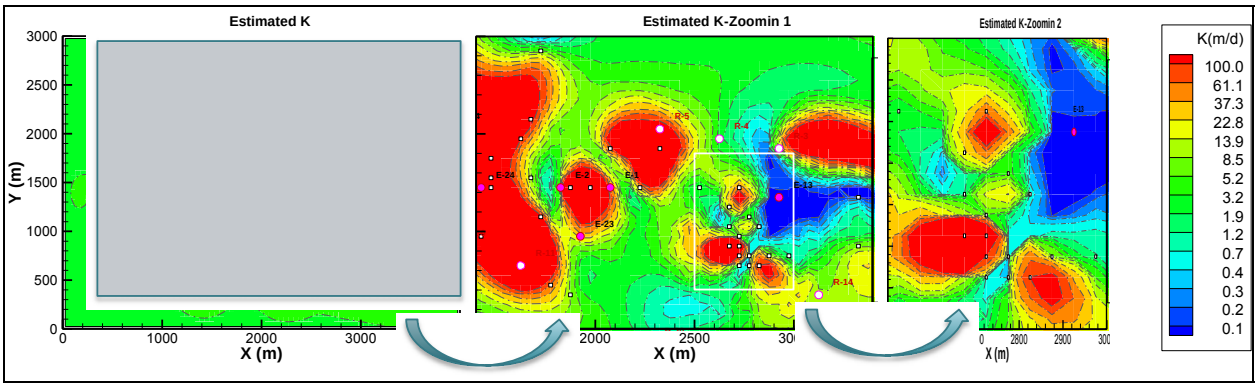


Figure 6-9. Delineated K-field at Different Resolution

The 3-D HT results suggest that there is a thin low K zone (notice that vertical and horizontal axes are not to the scale) in the middle of the aquifer (50 ~80 m) and it fades to the north and west directions. This is consistent with the geological description of the pinchout. Moreover, the top low K zone is also consistent with the geological description. Although this comparison is provided only in the qualitative sense, the consistency of geological information and the estimates enhances the credibility of the inverse results.

6.4 PERFORMANCE OBJECTIVE: ILLUSTRATE COST-EFFECTIVENESS OF HT AGAINST CONVENTIONAL TECHNIQUES

The cost-effectiveness of HT depends strongly on whether new wells and water treatment system are needed. The costs of performing pumping tests in the HT survey are the same as the costs to perform conventional pumping tests. The labor costs for performing HT analysis using all HT survey data are similar to the labor costs of performing conventional data analysis.

For large models, more expensive computer systems might be needed if computational time is a concern and more powerful computers are desirable. Using the same existing well network and pump-and-treat system, HT provides hydraulic information in greater details and higher resolution than conventional methods. To obtain similar detail levels using conventional methods, more wells and/or other local-scale measurement, such as cone penetrometer test, will be needed.

For the demonstration at the NCRS, an existing well network was available. The extracted groundwater was clean and no treatment was needed. The disposal of water to the existing storm drain was free. A high-performance computer system was available. The only costs for conducting the HT site characterization were the labor costs for conducting the pumping tests and performing the model inversion.

For the demonstration at the AFP44 site, the existing well network and pump-and-treat system were used. The available high-performance computer system was used to perform the HT inversion.

The only costs for conducting the HT site characterization were the labor costs for conducting the pumping tests and performing the HT model inversion.

Although we visited both sites frequently to download the transducer data to perform interim HT model inversion for this project, this amount of activity might not be necessary for other projects if the reduction of expenses is desired. Transducers can be programmed to record all data until multiple pumping tests are completed, and all the collected data can be analyzed at once to produce a single final model.

6.5 PERFORMANCE OBJECTIVE: ILLUSTRATE THAT HT IS 'USER-FRIENDLY'

HT is a "user-friendly" site characterization technology. HT surveys involve installing new wells, if needed, and performing pumping tests. The skills and equipment needed are the same as those commonly used in conventional site characterization. HT analysis involves compilation of pumping test data and performing model inversion. The input data required for model inversion are the same as the data used in groundwater model development and calibration, such as the input data for parameter estimation using the commonly used software PEST and MODFLOW.

6.6 PERFORMANCE OBJECTIVE: ILLUSTRATE THAT HT IS ABLE TO IDENTIFY LOW-CONDUCTIVITY ZONES

The evaluation described in Section 6.3 has illustrated that HT is able to delineate low-K zones consistent with the available lithologic data locally. In addition, it can infer the hydraulic continuity of the low-K regimes in between available lithologic information. It provides information as to whether these regimes are hydraulically functioning as competent barriers. In conjunction with available chemical concentration data, the information is useful for evaluating potential residual sources.

7.0 COST ASSESSMENT

There are two components constituting the total costs of HT site characterization. One component is the costs of conducting HT surveys in the field. This includes the costs associated with preparing and performing the field activities for collecting drawdown data from pumping tests. The second component is the costs of analyzing the data collected and interpreting the results.

Figure 7-1 conceptually illustrates the logistics of the HT Investigation planning process. The total cost depends on the desirable spatial resolution of the K-field to address the site characterization objectives; whether the existing well network is adequate for monitoring the hydraulic responses to the HT pumping tests; site access and operational constraints; whether onsite treatment system and disposal can be utilized; and the amount and noisiness of the data collected for HT model inversion.

Site Characterization Objectives

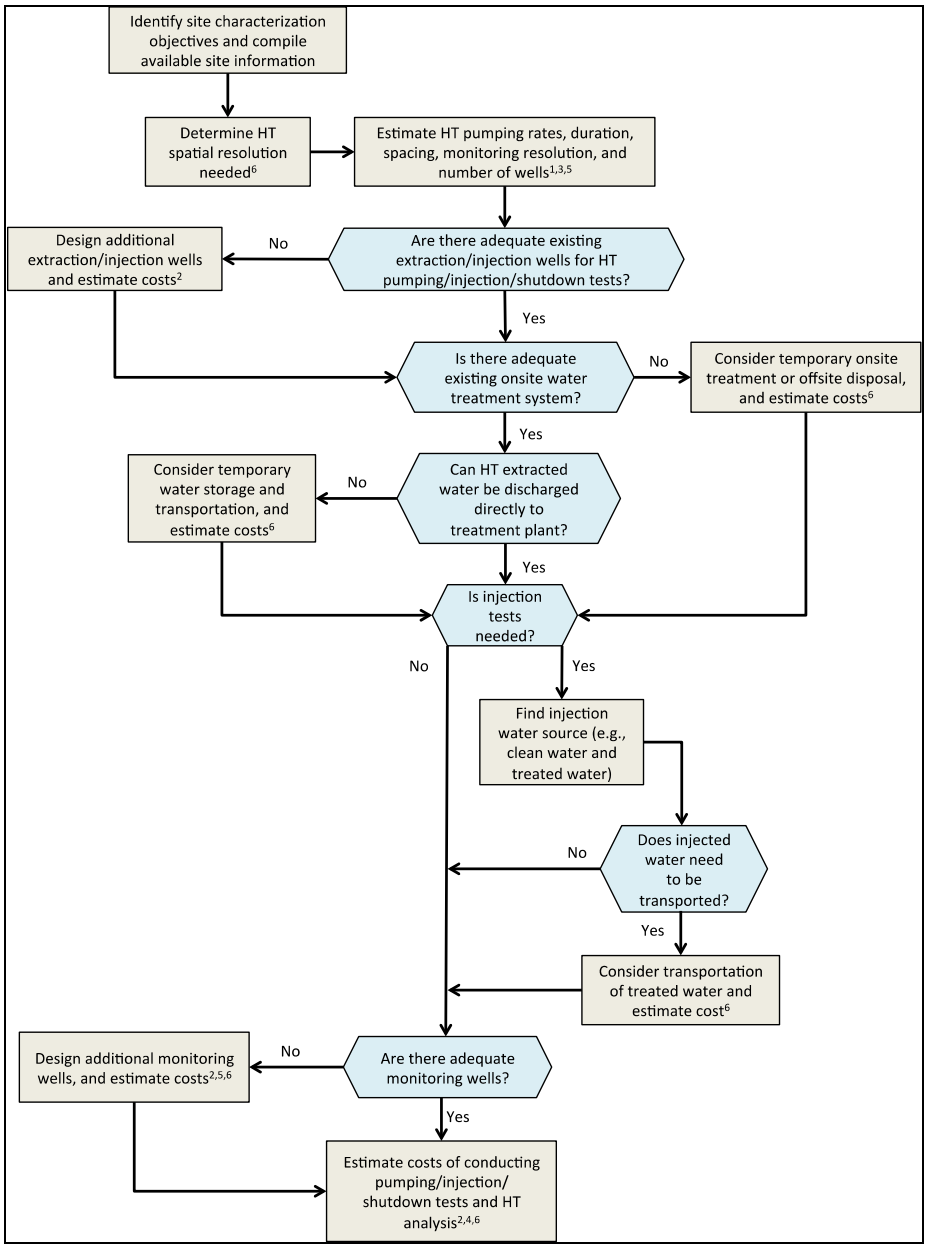
The specific objectives of a site characterization using HT and the situation at the site dictate the appropriate level of investigation efforts needed and the associated costs. For example, the extent and spatial resolution of the HT investigation for characterizing a paleochannel to support the design of a pump-and-treat containment system would be different from those for delineating pathways to support substrate delivery for enhancing source zone bioremediation.

Spatial Resolution

The desirable spatial resolution of the K- and S_s-fields to be delineated by HT depends on the site characterization objectives and the level of heterogeneity at the site. The number of extraction/injection wells and their pumping rates should be sufficient to generate hydraulic stresses that can be detected with adequate accuracy throughout the area of interest. Well locations should be appropriately selected to reduce costs. The number of monitoring wells, their locations and screen intervals, should be selected to cover the area of interest, and the spacing between wells should be smaller than the desirable length scale of the K- and S_s-fields to be delineated. The more non-redundant hydraulic response data is collected, the smaller will be the uncertainty and the greater will be the reliability of the HT results.

Existing Well Network

HT investigation is the most cost-effective if the existing well network at a site is sufficient, and there is no need to install additional wells. Wells screened in specific short depth intervals would be more preferable. If only long-screened wells are available, it is desirable to select wells with depth-discrete information, such as borehole flowmeter profiles and geophysical logs. If applicable, packers or multilevel liners can be installed to target specific hydrogeologic zones. If additional wells are needed to supplement the existing well network, multilevel/multichannel wells or clustered wells should be considered, as they would provide higher resolution results.



1. *Yeh, T.-C. J. and S. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, Water Resources Research, Vol. 36, No. 8, Pages 2095-2105.*
2. *Illman, W.A., S.J. Berg, and M. Alexander (2012), Cost comparisons of aquifer heterogeneity characterization methods, Ground Water Monitoring and Remediation, Vol. 32, No. 2, Pages 57-65.*
3. *Sun, R., T.-C.J. Yeh, D. Mao, M. Jin, W. Lu, and Y. Hao (2013), A temporal sampling strategy for hydraulic tomography analysis, Water Resources Research, Vol. 49, No. 1, Pages 1-16.*
4. *Berg, S.J. and W.A. Illman (2013), Field study of subsurface heterogeneity with steady-state hydraulic tomography, Ground Water, Vol. 51, No. 1, Pages 29-40.*
5. *Illman, W.A. (2014), Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks, Ground Water, Vol 52, No. 5, Pages 659-684.*
6. *Mok, C.M., T.-C.J. Yeh, and W.A. Illman (in preparation), Practical Implementation of Hydraulic Tomography.*

Figure 7-1. Conceptual Illustration of the Logistics of HT Planning Process

Site Access and Operational Constraints

Site access and operational constraints affect HT data collection. More pressure transducers with larger datalogging memory size can be used to reduce the need for site access and operational interference.

Onsite Treatment and Disposal

The groundwater extracted during HT investigation in chemically impacted zones would require treatment before disposal or re-injection into the subsurface. HT investigation could be more cost-effective if water from extraction wells are directly piped to the onsite treatment. If the onsite treatment unit has sufficient capacity and is available, but direct piping is not feasible, the extracted groundwater needs to be transported to the unit. Temporary storage units, such as water tanks, might be needed. If an onsite treatment unit is unavailable, a temporary treatment unit or off-site disposal might be needed. Alternatives, injection of clean water can be used to induce hydraulic stress. Regulatory requirements may apply in jurisdicted basins where laboratory analysis of water samples is needed to confirm that the injected water meets the water quality criteria.

7.1 COST MODEL

Table 7-1 summarizes the key cost components of conducting the HT site characterization. The first three items are related to conducting HT surveys. The last item is the cost of performing HT analysis.

Table 7-1. Cost Model for HT Site Characterization

Cost Element	Cost Components	Data Tracked During the Demonstration
Installation of Extraction/Injection Wells and Monitoring Wells	Unit: $ per linear foot of well Data requirements: • Number of wells, their diameters, depths, and screen intervals • Recommended installation method • Mobilization cost • Time required, Personnel required, and associated labor • Materials	NA (existing wells were used)
Groundwater Extraction, Treatment, and Disposal	Unit: • $ per pump • $ per volume of groundwater extracted • $ per operation day Data requirements: • Number of pumping tests, pumping rates, lift, and duration • Groundwater storage method • Groundwater treatment and disposal method • Time required, Personnel required, and associated labor • Materials • Analytical laboratory costs	NA (existing pump-and-treat system was used at AFP44; no need to treat extracted groundwater at NCRS)
Pumping Tests	Unit: • $ per day Data requirements: • Number of pumping tests, number of wells to be monitored, and duration • Pressure transducers • Time required, Personnel required, and associated labor • Materials	• Pump rates over time • Hydraulic head over time • Atmospheric pressure over time (for correction of hydraulic head)
Compilation of Pumping Test Data and HT Model Inversion	Unit: • $ per day Data requirements: • Number of transducers • Resolution of HT inversion model	NA (the level of details and experimentation of different analysis approaches was more involved than normal application)

7.1.1 Installation of Extraction/Injection Wells and Monitoring Wells

This cost is applicable only if the existing well network is inadequate. The cost depends on the following factors:

- The number of wells needed to provide a spatial resolution appropriate to meet the site characterization objectives;
- The depths and screen intervals needed to cover the depth range of concern and the vertical resolution appropriate in regard to characterization needs;
- The size of the wells to support the pumping rates needed to produce measurable hydraulic responses at the monitoring wells;
- Ease of site access;
- Subcontractor cost for installing the wells;
- Labor for overseeing the installation of the wells and obtaining the necessary permits; and
- The materials used for well installation.

7.1.2 Groundwater Extraction, Treatment, and Disposal

This cost is applicable only if the existing pump-and-treat system is inadequate. The cost depends on the following factors:

- The size of pumps must support the pumping rates and lift needed to produce measurable hydraulic responses at the monitoring wells;
- The size and type of treatment units supporting the volume and rates of groundwater extraction;
- The size of storage units to accommodate the extracted groundwater;
- The consumables for treating the extracted groundwater;
- Labor and materials for installing the equipment;
- Necessary analytical laboratory testing;
- Labor for acquiring the necessary permits.

7.1.3 Conducting Pumping Tests

The costs of conducting the HT surveys include the costs of labor required to prepare and perform the sequential pumping tests. Equipment costs include pressure transducers and data acquisition system. The labor costs typically depend upon the type of personnel conducting the surveys and their associated labor rates and hours.

The total cost for multiple tomography surveys depends on the cost per single survey. Since the labor hours depend on the number of pumping locations utilized and the number of observation ports to be monitored, the cost per survey is site specific. In addition, since the equipment can be rented for the HT survey at a site, the equivalent equipment rental costs can be considered for inclusion.

7.1.4 Compilation of Pumping Test Data and HT Model Inversion

This cost element covers the labor, software, and equipment costs associated with processing the tomography data using a groundwater flow model and inverse modeling algorithms to produce tomograms of K and Ss parameters. It also includes the labor costs for interpreting the results, which depends on the quality of the data, whether significant effort is needed to remove noise, trend, and fluctuations. It also depends on the number of transducers, resolution and the size of the model.

7.2 COST DRIVERS

Table 7-2 summarizes the cost items and the potential cost drivers for conducting HT site characterization. Installing new wells and treating the extracted groundwater are the major cost drivers. HT site characterization is most cost-effective when it is performed using an existing well network and treatment system. HT can always be readily applicable to sites with existing pump-and-treat systems, since it optimally uses all the information available. As such, the results are the most optimal and unbiased, based on the available information. The resolution is dictated by the spacing of the existing well network.

Table 7-2. Cost Items and Cost Drivers for HT Site Characterization

Cost Items	Cost Factors	Remarks
Extraction/injection well network: wells and pumps	Availability of existing wells; Number of new wells, their well sizes and depths; Ease of access; Permitting; Pump size and packers, if needed	Potential cost driver if new wells are needed
Monitoring well network	Availability of existing wells; Number of new wells and depths; Ease of access; Permitting	Expensive, but less costly than extraction wells.
Extracted water disposal	Availability of on-site treatment; extraction rate and duration; storage and treatment costs; transportation costs if applicable.	Potentially expensive
Transducers	Number of transducers; Type, size, and storage	Relatively inexpensive if diameter of well/sounding tube greater than or equal to 2."
Sequential HT aquifer tests	Availability of site staff	Relatively inexpensive if site staff is available
HT data analysis	Resolution needed; 2D versus 3D; steady state versus transient	Relatively inexpensive

7.3 COST ANALYSIS

For a cost comparison, we use a project of similar scale to the NCRS. Here, we considered two approaches for the characterization of subsurface heterogeneity: one that relies on detailed borehole data (Approach 1) and another that relies on the inverse modeling of several pumping tests (Approach 2). Where possible, real costs from this study are used, and labor is estimated at approximately $100/hr in order to better reflect the costs of the environmental industry. Table 7-3 and Table 7-4 summarize the costs for Approaches 1 and 2, respectively.

For Approach 1, we assumed that continuous soils cores are collected during the installation of fully screened wells. This is a slow process, which can add considerable costs to drilling in comparison to the traditional installation of pumping or observation wells. The drilling costs for the five wells totaled $60,000. However, if coring is the only objective, and if wells do not need to be installed, then the costs can be somewhat reduced. Another costly item for Approach 1 is the laboratory permeameter analysis of soil cores, which can be a very slow process when a large number of samples need to be analyzed and when the analysis is undertaken for lower K materials. For our study, the estimated duration for sample analysis is based on the experience of Alexander (2009) who performed the laboratory analysis at the University of Waterloo. The cost of sample analysis amounted to $100 per sample, or $47,100 for all 471 samples. The data analysis component included data reduction/processing ($4,000), geostatistical analysis ($4,000), and report writing ($2,000). The total cost of characterizing the subsurface heterogeneity through the geostatistical analysis of core samples (Approach 1) was $117,100.

For Approach 2, we separated the costs associated with the hydraulic tomography survey and the required equipment that may be reused. As in Approach 1, drilling for Approach 2 is a slow process because of the installation of multi-screen wells. Here, drilling is assumed to be complete without the collection of soil cores. A significant amount of cost is added to the drilling by alternating the backfilling of sand pack and bentonite in order to prevent short-circuiting between adjacent pumping and/or observation intervals. The total estimated cost for drilling is approximately $60,000 and is based on the cost of installing all five wells.

Costs to conduct hydraulic tomography include the man hours required to perform multiple pumping tests ($12,000), data processing ($8,000), the inverse modeling of the test results ($12,000), and reporting ($2,000). The costs of conducting hydraulic tomography minus the equipment costs resulted in $94,000. Therefore, we see that Approaches 1 and 2 are comparable in terms of costs if the equipment costs for hydraulic tomography are not accounted for.

Equipment costs include the pressure transducers ($30,000), the CMT systems ($5,000), FLUTe liners with five pressure transducers each ($36,000), a double-packer system with a submersible pump ($5,000), a data acquisition system ($8,000), and a high-end workstation or a PC-cluster ($20,000). The equipment costs add up to $104,000; however, many of these items can be reused, except for the CMT system, which we assume will remain at the site upon completion of the survey.

Table 7-3. Cost Estimate for Heterogeneity Characterization Relying on Point Data

Detailed Characterization	Estimated Costs
1. Drilling (with complete core collection)	$60,000
2. Permeameter analysis (471 samples @ 1sample/hour)	$47,100
3. Data analysis	
Data Processing (1 week)	$4,000
Geostatistical analysis (1 week)	$4,000
Reporting (0.5 weeks)	$2,000
Total (1+2+3)	$117,000

Table 7-4. Cost Estimate for Performing HT

Transient HT	**Estimated Costs**
1. Drilling (with complete core collection)	$60,000
2. Conducting 4 x 24 hours pumping tests	$12,100
3. Data analysis	
Data Processing (2 weeks)	$8,000
Inverse modeling (3 weeks)	$12,000
Reporting (0.5 week)	$2,000
4. Subtotal (1+2+3)	$94,000
Capital Costs	**Estimated Costs**
5. Instrumentation	
Pressure transducers (28 CMT, 6 for 2" wells)	$30,000
CMT systems	$5,000
FLUTe liners (with five transducers each)	$36,000
Pump-Packer system	$5,000
Data acquisition system	$8,000
6. PC cluster for modeling	$20,000
7. Subtotal capital costs (5+6)	$104,000
Total (4+7)	**$198,000**

While these estimates are very approximate, they do suggest that implementing hydraulic tomography can be cost-effective if one considers the equipment as a separate cost item, some of which can be reused in other projects. For example, the instrumentation required for conducting the pumping tests and monitoring drawdown can be used at other sites. The same can be said for the computer cluster used for running the inverse model. Once purchased, this cluster can be used for many sites, and the costs can be spread out over many applications. Most importantly, it has been demonstrated that HT significantly improved predictions of drawdowns when compared to conventional methods. The reliance on pumping test data using hydraulic tomography, as opposed to permeameter data, may also be another reason why hydraulic tomography performed better than the geostatistics approach, as small scale samples can be disturbed and core recovery is not always complete.

Regardless of the choice in characterization method, we contend that improved site characterization before implementing remediation systems will lead to more efficient and effective clean-up operations. Thus, the costs spent upfront to accurately characterize the site should minimize issues that could arise later due to poor site characterization.

8.0 IMPLEMENTATION ISSUES

8.1 HT INVESTIGATION PLANNING

Figure 7-1 shows a flowchart conceptually illustrating the logistics of the planning process for HT Investigation. Some of these items are discussed in the following sections and consider specific site characterization objectives and the site situation.

After setting the required pumping well spacing, the number and location of potential new pumping wells have to be determined. The handling of the extracted water has to be accounted for. If the on-site water treatment system is not available or not suitable for the extracted water, temporary storage and transportation options should be discussed, with consideration of the pumping rates and durations required for showing sufficient drawdown responses. If injection tests are required for the characterization, a suitable source of injection water, such as clean or treated water, needs to be found and its transportation planned accordingly. Depending on the spacing of the existing monitoring well network, the number and location of new monitoring wells also have to be determined.

8.2 POTENTIAL REGULATIONS

If additional wells are needed, and especially if they need to be installed in areas with high chemical concentration, pertinent regulatory approval and permits might be required. This is a similar issue with conventional well installation.

If the HT pumping tests involve groundwater extraction, pumping permits might be required. In addition, permits for the discharge to the on-site or off-site treatment systems need to be acquired. Depending on the application process, extraction water sampling might be necessary. Similarly, permits might have to be obtained for the water injection, with a potential sampling of the injection water.

8.3 CONCERNS, RESERVATIONS, AND DECISION-MAKING FACTORS

The key factors to be considered in making a decision on whether HT is appropriate for a site include cost-effectiveness, timing and duration, knowledge of background hydraulic stresses, and chemical mobilization. The cost-effectiveness depends on the appropriate number of wells, which is dictated by the spatial resolution needed to meet the objectives and whether existing wells and treatment system are adequate. If existing well and treatment system can be utilized, the costs associated with HT is minimal. Since HT relies primarily on hydraulic response data, it is best applicable to sites with known background hydraulic information, such as presence of other pumping wells and water-level fluctuations. In addition, water level changes due to HT pumping tests might cause chemicals to move during the tests. The duration of the pumping tests are usually short, and the amount of the associated chemical movement is typically small. However, if the aquifer is very permeable, a large pumping rate might be required to generate a measurable hydraulic response signal. On the other hand, if the aquifer is relatively impermeable, the well yield might be small, and a longer HT pumping test duration might be needed.

8.4 RELEVANT PROCUREMENT ISSUES

Standard commercial equipment for groundwater extraction, injection, and monitoring, such as pumps, monitoring wells, liners, packers, and pressure transducers, is suitable for HT. For HT analysis, adequate computational power is needed, possibly in the form of computing clusters.

9.0 REFERENCES

AECOM. 2010. Installation restoration program environmental remediation annual update sites 14 and 17 July through June 2009. Tucson International Airport Area. Tucson, Arizona. Denver, Colorado.

AECOM. 2011. Installation restoration program environmental remediation annual update--site OT-12 July 2010 through June 2011. Tucson International Airport Area, Tucson, Arizona. Denver, Colorado.

AECOM. 2012. Volume III Conceptual site model, in Focused remedial investigation/feasibility study for 1,4-dioxane: Tucson International Airport Area Superfund Site Area A, Tucson, Arizona. Denver, Colorado.

Alexander, M. 2009. "Evaluation of Traditional Hydrogeologic Characterization Approaches in a Highly Heterogeneous Glaciofluvial Aquifer/Aquitard System." UWSpace.

Alexander, M., Berg, S. J., and Illman, W. A. (2011), Field Study of Hydrogeologic Characterization Methods in a Heterogeneous Aquifer, *Ground Water* 49 (3):365-382. doi: 10.1111/j.1745-6584.2010.00729.x.

Berg, S. J., and Illman, W. A. (2011), Capturing aquifer heterogeneity: Comparison of approaches through controlled sandbox experiments, *Water Resources Research* 47 (9):n/a-n/a. doi: 10.1029/2011WR010429.

Berg, S. J., and Illman, W. A. (2011), Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system, *Water Resources Research* 47 (10):n/a-n/a. doi: 10.1029/2011WR010616.

Berg, S. J., and Illman, W. A. (2013), Field Study of Subsurface Heterogeneity with Steady-State Hydraulic Tomography, *Ground Water* 51 (1):29-40. doi: 10.1111/j.1745-6584.2012.00914.x.

Berg, S. J., and Illman, W. A. (2015), Comparison of Hydraulic Tomography with Traditional Methods at a Highly Heterogeneous Site, *Groundwater* 53 (1):71-89. doi: 10.1111/gwat.12159.

Bohling, G. C., and Butler, J. J. (2010), Inherent Limitations of Hydraulic Tomography, *Ground Water* 48 (6):809-824. doi: 10.1111/j.1745-6584.2010.00757.x.

Bohling, G. C., Butler, J. J., Zhan, X., and Knoll, M. D. (2007), A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities, *Water Resources Research* 43 (5):n/a-n/a. doi: 10.1029/2006WR004932.

Bohling, G. C., Zhan, X., Butler, J. J., and Zheng, L. (2002), Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities, *Water Resources Research* 38 (12):60-1-60-15. doi: 10.1029/2001WR001176.

Brauchler, R., Cheng, J. T., Dietrich, P., Everett, M., Johnson, B., Liedl, R., and Sauter, M. (2007), An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion, *Journal of Hydrology* 345 (3–4):184-198. doi: https://doi.org/10.1016/j.jhydrol.2007.08.011.

Brauchler, R., Doetsch, J., Dietrich, P., and Sauter, M. (2012), Derivation of site-specific relationships between hydraulic parameters and p-wave velocities based on hydraulic and seismic tomography, *Water Resources Research* 48 (3):n/a-n/a. doi: 10.1029/2011WR010868.

Brauchler, R., Hu, R., Dietrich, P., and Sauter, M. (2011), A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography, *Water Resources Research* 47 (3):n/a-n/a. doi: 10.1029/2010WR009635.

Brauchler, R., Hu, R., Hu, L., Jiménez, S., Bayer, P., Dietrich, P., and Ptak, T. (2013), Rapid field application of hydraulic tomography for resolving aquifer heterogeneity in unconsolidated sediments, *Water Resources Research* 49 (4):2013-2024. doi: 10.1002/wrcr.20181.

Brauchler, R., Liedl, R., and Dietrich, P. (2003), A travel time based hydraulic tomographic approach, *Water Resources Research* 39 (12):n/a-n/a. doi: 10.1029/2003WR002262.

Cardiff, M., and Barrash, W. (2011), 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response, *Water Resources Research* 47 (12):n/a-n/a. doi: 10.1029/2010WR010367.

Cardiff, M., Barrash, W., and Kitanidis, P. K. (2012), A field proof-of-concept of aquifer imaging using 3-D transient hydraulic tomography with modular, temporarily-emplaced equipment, *Water Resources Research* 48 (5):n/a-n/a. doi: 10.1029/2011WR011704.

Cardiff, M., Barrash, W., and Kitanidis, P. K. (2013), Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities, *Water Resources Research* 49 (11):7311-7326. doi: 10.1002/wrcr.20519.

Cardiff, M., Barrash, W., Kitanidis, P. K., Malama, B., Revil, A., Straface, S., and Rizzo, E. (2009), A Potential-Based Inversion of Unconfined Steady-State Hydraulic Tomography, *Ground Water* 47 (2):259-270. doi: 10.1111/j.1745-6584.2008.00541.x.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005), Inverse problem in hydrogeology, *Hydrogeology Journal* 13 (1):206-222. doi: 10.1007/s10040-004-0404-7.

Carrera, J., and Glorioso, L. (1991), On geostatistical formulations of the groundwater flow inverse problem, *Advances in Water Resources* 14 (5):273-283. doi: http://dx.doi.org/10.1016/0309-1708(91)90039-Q.

Castagna, M., and Bellin, A. (2009), A Bayesian approach for inversion of hydraulic tomographic data, *Water Resources Research* 45 (4):W04410. doi: 10.1029/2008WR007078.

Chen, Y., and Oliver, D. S. (2013), Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification, *Computational Geosciences* 17 (4):689-703. doi: 10.1007/s10596-013-9351-5.

Chen, G., Illman, W. A., Thompson, D. L., Vesselinov, V. V., and Neuman, S. P. (2000), Geostatistical, Type-Curve, and Inverse Analyses of Pneumatic Injection Tests in Unsaturated Fractured Tuffs at the Apache Leap Research Site Near Superior, Arizona, *Dynamics of Fluids in Fractured Rock*:73-98.

Earth Tech. 1992. Remedial Investigation Report, Prepared for Air Force Plant 44, Tucson Arizona. Tempe Arizona.

Earth Tech. 2007. Summary of Reclamation Well Field and Soil Remediation Operations Annual Update July 2006 through June 2007 US AFP 44. Englewood, Colorado.

Fienen, M., Hunt, R., Krabbenhoft, D., and Clemo, T. (2009), Obtaining parsimonious hydraulic conductivity fields using head and transport observations: A Bayesian geostatistical parameter estimation approach, *Water Resources Research* 45 (8):n/a-n/a. doi: 10.1029/2008WR007431.

Hanson, R. T., and Benedict, J. F. 1994. Simulation of ground-water flow and potential land subsidence, upper Santa Cruz Basin, Arizona. In *Water-Resources Investigations Report*.

Houser, B. B. P., L.; Esser, R. P.; Gettings, M. E. 2004. Stratigraphic and tectonic history of the Tucson Basin, Pima county, Arizona, based on the Exxon State (32)-1 well, U.S. Geological Survey Scientific Investigations Report 2004-5076. Tucson, Arizona.

Hsieh, P. A., Neuman, S. P., Stiles, G. K., and Simpson, E. S. (1985), Field Determination of the Three-Dimensional Hydraulic Conductivity Tensor of Anisotropic Media: 2. Methodology and Application to Fractured Rocks, *Water Resources Research* 21 (11):1667-1676. doi: 10.1029/WR021i011p01667.

Hu, R., Brauchler, R., Herold, M., and Bayer, P. (2011), Hydraulic tomography analog outcrop study: Combining travel time and steady shape inversion, *Journal of Hydrology* 409 (1–2):350-362. doi: https://doi.org/10.1016/j.jhydrol.2011.08.031.

Huang, S.-Y., Wen, J.-C., Yeh, T.-C. J., Lu, W., Juan, H.-L., Tseng, C.-M., Lee, J.-H., and Chang, K.-C. (2011), Robustness of joint interpretation of sequential pumping tests: Numerical and field experiments, *Water Resources Research* 47 (10):n/a-n/a. doi: 10.1029/2011WR010698.

HydroGeoLogic. 2012. Final comprehensive well survey report, Air Force Plant 44, Tucson, Arizona. Phoenix, Arizona.

Illman, W. A. (2014), Hydraulic Tomography Offers Improved Imaging of Heterogeneity in Fractured Rocks, *Groundwater* 52 (5):659-684. doi: 10.1111/gwat.12119.

Illman, W. A., Berg, S. J., and Alexander, M. (2012), Cost Comparisons of Aquifer Heterogeneity Characterization Methods, *Ground Water Monitoring & Remediation* 32 (2):57-65. doi: 10.1111/j.1745-6592.2011.01376.x.

Illman, W. A., Berg, S. J., and Zhao, Z. (2015), Should hydraulic tomography data be interpreted using geostatistical inverse modeling? A laboratory sandbox investigation, *Water Resources Research* 51 (5):3219-3237. doi: 10.1002/2014WR016552.

Illman, W. A., Craig, A. J., and Liu, X. (2008), Practical Issues in Imaging Hydraulic Conductivity through Hydraulic Tomography, *Ground Water* 46 (1):120-132. doi: 10.1111/j.1745-6584.2007.00374.x.

Illman, W. A., Liu, X., and Craig, A. (2007), Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms, *Journal of Hydrology* 341 (3–4):222-234. doi: https://doi.org/10.1016/j.jhydrol.2007.05.011.

Illman, W. A., Liu, X., Takeuchi, S., Yeh, T.-C. J., Ando, K., and Saegusa, H. (2009), Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan, *Water Resources Research* 45 (1):n/a-n/a. doi: 10.1029/2007WR006715.

Illman, W. A., and Neuman, S. P. (2001), Type curve interpretation of a cross-hole pneumatic injection test in unsaturated fractured tuff, *Water Resources Research* 37 (3):583-603. doi: 10.1029/2000WR900273.

Illman, W. A., and Neuman, S. P. (2003), Steady-state analysis of cross-hole pneumatic injection tests in unsaturated fractured tuff, *Journal of Hydrology* 281 (1–2):36-54. doi: https://doi.org/10.1016/S0022-1694(03)00199-9.

Illman, W. A., Zhu, J., Craig, A. J., and Yin, D. (2010), Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study, *Water Resources Research* 46 (4):n/a-n/a. doi: 10.1029/2009WR007745.

Karrow, P. F. (1993), Quaternary geology, Stratford-Conestogo area: [Toronto] : Ontario Ministry of Northern Development and Mines, 1993.

Koltermann, C. E., and Gorelick, S. M. (1996), Heterogeneity in Sedimentary Deposits: A Review of Structure-Imitating, Process-Imitating, and Descriptive Approaches, *Water Resources Research* 32 (9):2617-2658. doi: 10.1029/96WR00025.

Konikow, L. F., and Bredehoeft, J. D. (1992), Ground-water models cannot be validated, *Advances in Water Resources* 15 (1):75-83. doi: http://dx.doi.org/10.1016/0309-1708(92)90033-X.

Lavenue, M., and de Marsily, G. (2001), Three-dimensional interference test interpretation in a fractured aquifer using the Pilot Point Inverse Method, *Water Resources Research* 37 (11):2659-2675. doi: 10.1029/2000WR000289.

Leake, S., and Hanson, R. (1987), Distribution and movement of trichloroethylene in ground water in the Tucson area, Arizona. Vol. 86: US Geological Survey.

Lee, J., and Kitanidis, P. K. (2014), Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA), *Water Resources Research* 50 (7):5410-5427. doi: 10.1002/2014WR015483.

Li, W., Englert, A., Cirpka, O. A., Vanderborght, J., and Vereecken, H. (2007), Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests, *Water Resources Research* 43 (4):n/a-n/a. doi: 10.1029/2006WR005333.

Li, W., Englert, A., Cirpka, O. A., and Vereecken, H. (2008), Three-Dimensional Geostatistical Inversion of Flowmeter and Pumping Test Data, *Ground Water* 46 (2):193-201. doi: 10.1111/j.1745-6584.2007.00419.x.

Li, W., Nowak, W., and Cirpka, O. A. (2005), Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown, *Water Resources Research* 41 (8):n/a-n/a. doi: 10.1029/2004WR003874.

Liu, X., Illman, W. A., Craig, A. J., Zhu, J., and Yeh, T. C. J. (2007), Laboratory sandbox validation of transient hydraulic tomography, *Water Resources Research* 43 (5):n/a-n/a. doi: 10.1029/2006WR005144.

Liu, X., Zhou, Q., Kitanidis, P. K., and Birkholzer, J. T. (2014), Fast iterative implementation of large-scale nonlinear geostatistical inverse modeling, *Water Resources Research* 50 (1):198-207. doi: 10.1002/2012WR013241.

Mao, D., Wan, L., Yeh, T.-C. J., Lee, C.-H., Hsu, K.-C., Wen, J.-C., and Lu, W. (2011), A revisit of drawdown behavior during pumping in unconfined aquifers, *Water Resources Research* 47 (5):n/a-n/a. doi: 10.1029/2010WR009326.

Mao, D., Yeh, T.-C. J., Wan, L., Wen, J.-C., Lu, W., Lee, C.-H., and Hsu, K.-C. (2013a), Joint interpretation of sequential pumping tests in unconfined aquifers, *Water Resources Research* 49 (4):1782-1796. doi: 10.1002/wrcr.20129.

Mao, D., Yeh, T.-C. J., Wan, L., Lee, C.-H., Hsu, K.-C., Wen, J.-C., and Lu, W. (2013b), Cross-correlation analysis and information content of observed heads during pumping in unconfined aquifers, *Water Resources Research* 49 (2):713-731. doi: 10.1002/wrcr.20066.

Martin, P. J., and Frind, E. G. (1998), Modeling a Complex Multi-Aquifer System: The Waterloo Moraine, *Ground Water* 36 (4):679-690. doi: 10.1111/j.1745-6584.1998.tb02843.x.

McLaughlin, D., and Townley, L. R. (1996), A Reassessment of the Groundwater Inverse Problem, *Water Resources Research* 32 (5):1131-1161. doi: 10.1029/96WR00160.

Meyer, P. A. Browers., M.; and Martin, P. J. (2014), A three-dimensional groundwater flow model of the Waterloo Moraine for water resource management, *Canadian Water Resources Journal / Revue canadienne des ressources hydriques* 39 (2):167-180. doi: 10.1080/07011784.2014.914800.

Montgomery and Associates. 2015. Installation of Nine Upper Zone Groundwater Monitor Wells, U.S. Air Force Plant 44, Tucson, Arizona. Tucson, Arizona.

Neuman, S. P. (1975), Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response, *Water Resources Research* 11 (2):329-342. doi: 10.1029/WR011i002p00329.

Neuman, S. P., Walter, G. R., Bentley, H. W., Ward, J. J., and Gonzalez, D. D. (1984), Determination of Horizontal Aquifer Anisotropy with Three Wells, *Ground Water* 22 (1):66-72. doi: 10.1111/j.1745-6584.1984.tb01477.x.

Refsgaard, J. C., Christensen, S., Sonnenborg, T. O., Seifert, D., Højberg, A. L., and Troldborg, L. (2012), Review of strategies for handling geological uncertainty in groundwater flow and transport modeling, *Advances in Water Resources* 36:36-50. doi: http://dx.doi.org/10.1016/j.advwatres.2011.04.006.

Schöniger, A., Illman, W. A., Wöhling, T., and Nowak, W. (2015), Finding the right balance between groundwater model complexity and experimental effort via Bayesian model selection, *Journal of Hydrology* 531:96-110. doi: http://dx.doi.org/10.1016/j.jhydrol.2015.07.047.

Schöniger, A., Nowak, W., and Hendricks Franssen, H. J. (2012), Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography, *Water Resources Research* 48 (4):n/a-n/a. doi: 10.1029/2011WR010462.

Sebol, L. A. 2000. "Determination of groundwater age using CFCs in three shallow aquifers in Southern Ontario." University of Waterloo.

Sharmeen, R., Illman, W. A., Berg, S. J., Yeh, T.-C. J., Park, Y.-J., Sudicky, E. A., and Ando, K. (2012), Transient hydraulic tomography in a fractured dolostone: Laboratory rock block experiments, *Water Resources Research* 48 (10):n/a-n/a. doi: 10.1029/2012WR012216.

Soueid Ahmed, A., Jardani, A., Revil, A., and Dupont, J. P. (2014), Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data, *Water Resources Research* 50 (4):3502-3522. doi: 10.1002/2013WR014645.

Soueid Ahmed, A., Zhou, J., Jardani, A., Revil, A., and Dupont, J. P. (2015), Image-guided inversion in steady-state hydraulic tomography, *Advances in Water Resources* 82:83-97. doi: https://doi.org/10.1016/j.advwatres.2015.04.001.

Straface, S., Yeh, T. C. J., Zhu, J., Troisi, S., and Lee, C. H. (2007), Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy, *Water Resources Research* 43 (7):n/a-n/a. doi: 10.1029/2006WR005287.

Sudicky, E. A., Unger, A. J. A., and Lacombe, S. (1995), A Noniterative Technique for the Direct Implementation of Well Bore Boundary Conditions in Three-Dimensional Heterogeneous Formations, *Water Resources Research* 31 (2):411-415. doi: 10.1029/94WR02854.

Sun, R., Yeh, T.-C. J., Mao, D., Jin, M., Lu, W., and Hao, Y. (2013), A temporal sampling strategy for hydraulic tomography analysis, *Water Resources Research* 49 (7):3881-3896. doi: 10.1002/wrcr.20337.

Tarantola, A. (2005), Inverse problem theory and methods for model parameter estimation: SIAM.

Tillman, F. D. 2009. Results of the analyses for 1,4-dioxane of groundwater samples collected in the Tucson Airport Remediation Project area, south-central Arizona, 2006-2009: U.S. Geological Survey Open-File Report 2009-1196.

Tso, M. C.-H., Zha, Y., Jim Yeh, T.-C., and Wen, J.-C. (2016), The relative importance of head, flux, and prior information in hydraulic tomography analysis, *Water Resources Research* 52 (1):3-20. doi: 10.1002/2015WR017191.

URS. 2013. Final treatability study work plan OT012--South of Los Reales Road regional groundwater plume Air Force Plant 44, Tucson, Arizona. Phoenix, Arizona.

URS. 2014. Annual optimized exit strategy performance metric report: July 2012 to June 2013 OT012-South of Los Reales Road regional groundwater plume Air Force Plant 44, Phoenix, Arizona.

Vasco, D., Keers, H., and Karasaki, K. (2000), Estimation of reservoir properties using transient pressure data: An asymptotic approach, *Water Resources Research* 36 (12):3447-3465.

Vesselinov, V. V., Neuman, S. P., and Illman, W. A. (2001a), Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 1. Methodology and borehole effects, *Water Resources Research* 37 (12):3001-3017. doi: 10.1029/2000WR000133.

Vesselinov, V. V., Neuman, S. P., and Illman, W. A. (2001b), Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 2. Equivalent parameters, high-resolution stochastic imaging and scale effects, *Water Resources Research* 37 (12):3019-3041. doi: 10.1029/2000WR000135.

Wen, J.-C., Wu, C.-M., Yeh, T.-C. J., and Tseng, C.-M. (2010), Estimation of effective aquifer hydraulic properties from an aquifer test with multi-well observations (Taiwan), *Hydrogeology Journal* 18 (5):1143-1155. doi: 10.1007/s10040-010-0577-1.

Wu, C.-M., Yeh, T.-C. J., Zhu, J., Lee, T. H., Hsu, N.-S., Chen, C.-H., and Sancho, A. F. (2005), Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resources Research* 41 (9):n/a-n/a. doi: 10.1029/2004WR003717.

Xiang, J., Yeh, T.-C. J., Lee, C.-H., Hsu, K.-C., and Wen, J.-C. (2009), A simultaneous successive linear estimator and a guide for hydraulic tomography analysis, *Water Resources Research* 45 (2):n/a-n/a. doi: 10.1029/2008WR007180.

Yeh, T. C. J., Khaleel, R., and Carroll, K. C. (2015), Flow Through Heterogeneous Geologic Media. New York, NY: Cambridge University Press.

Yeh, T. C. J., and Liu, S. (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resources Research* 36 (8):2095-2105. doi: 10.1029/2000WR900114.

Yeh, T. C. J., Mao, D., Wan, L., Lee, C. H., Wen, J. C., and Lu, W. (2012), Replies to comments on "A revisit of drawdown behavior during pumping in unconfined aquifers" by Neuman and Mishra, *Water Resources Research* 48 (2):n/a-n/a. doi: 10.1029/2011WR011153.

Yeh, T. C. J., Srivastava, R., Guzman, A., and Harter, T. (1993), A Numerical Model for Water Flow and Chemical Transport in Variably Saturated Porous Media, *Ground Water* 31 (4):634-644. doi: 10.1111/j.1745-6584.1993.tb00597.x.

Yeh, T. C. J., and Zhang, J. (1996), A Geostatistical Inverse Method for Variably Saturated Flow in the Vadose Zone, *Water Resources Research* 32 (9):2757-2766. doi: 10.1029/96WR01497.

Zha, Y., Yeh, T.-C. J., Illman, W. A., Onoe, H., Mok, C. M. W., Wen, J.-C., Huang, S.-Y., and Wang, W. (2017), Incorporating geologic information into hydraulic tomography: A general framework based on geostatistical approach, *Water Resources Research* 53 (4):2850-2876. doi: 10.1002/2016WR019185.

Zha, Y., Yeh, T.-C. J., Illman, W. A., Tanaka, T., Bruines, P., Onoe, H., and Saegusa, H. (2015), What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments, *Journal of Hydrology* 531:17-30. doi: http://dx.doi.org/10.1016/j.jhydrol.2015.06.013.

Zha, Y., Yeh, T.-C. J., Mao, D., Yang, J., and Lu, W. (2014), Usefulness of flux measurements during hydraulic tomographic survey for mapping hydraulic conductivity distribution in a fractured medium, *Advances in Water Resources* 71:162-176. doi: http://dx.doi.org/10.1016/j.advwatres.2014.06.008.

Zha, Y., Yeh, T.-C. J., Illman, W. A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., Mao, D., Takeuchi, S., and Wen, J.-C. (2016), An Application of Hydraulic Tomography to a Large-Scale Fractured Granite Site, Mizunami, Japan, *Groundwater* 54 (6):793-804. doi: 10.1111/gwat.12421.

Zha, Y., Yeh, T.-C. J., Shi, L., Huang, S.-Y., Wang, W., and Wen, J.-C. (2017), Quasi-steady state conditions in heterogeneous aquifers during pumping tests, *Advances in Water Resources*. doi: https://doi.org/10.1016/j.advwatres.2017.03.017.

Zhang, Z., and Brusseau, M. L. (1998), Characterizing Three-Dimensional Hydraulic Conductivity Distributions Using Qualitative and Quantitative Geologic Borehole Data: Application to a Field Site, *Ground Water* 36 (4):671-678. doi: 10.1111/j.1745-6584.1998.tb02842.x.

Zhang, Z., and Brusseau, M. L. (1999), Nonideal transport of reactive solutes in heterogeneous porous media: 5. Simulating regional-scale behavior of a trichloroethene plume during pump-and-treat remediation, *Water Resources Research* 35 (10):2921-2935. doi: 10.1029/1999WR900162.

Zhao, Z., and Illman, W. A. (2017), On the importance of geological data for three-dimensional steady-state hydraulic tomography analysis at a highly heterogeneous aquifer-aquitard system, *Journal of Hydrology* 544:640-657. doi: https://doi.org/10.1016/j.jhydrol.2016.12.004.

Zhao, Z., Illman, W. A., and Berg, S. J. (2016), On the importance of geological data for hydraulic tomography analysis: Laboratory sandbox study, *Journal of Hydrology* 542:156-171. doi: http://dx.doi.org/10.1016/j.jhydrol.2016.08.061.

Zhou, J., Revil, A., Karaoulis, M., Hale, D., Doetsch, J., and Cuttler, S. (2014), Image-guided inversion of electrical resistivity data, *Geophysical Journal International* 197 (1):292-309. doi: 10.1093/gji/ggu001.

Zhu, J., and Yeh, T.-C. J. (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resources Research* 41 (7):n/a-n/a. doi: 10.1029/2004WR003790.

APPENDIX A POINTS OF CONTACT

Point of Contact Name	Organization Name Address	Phone Fax Email	Role in Project
Dr. Chin Man W. Mok	GSI Environmental Inc.	510-316-8445 cmmok@gsienv.com	Principal Investigator
Dr. T.-C. Jim Yeh	University of Arizona	520-621-5943 yeh@hwr.arizona.edu	Co-Principal Investigator
Dr. Walter A. Illman	University of Waterloo	519-888-4567 x38341 willman@uwaterloo.ca	Co-Principal Investigator